GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION OF E. COLI

RELATED APPLICATIONS

This application claims priority from U.S. Provisional Patent Application Serial Number 60/173,005 filed December 23, 1999, the disclosure of which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

Since the discovery of penicillin, the use of antibiotics to treat the ravages of bacterial infections has saved millions of lives. With the advent of these "miracle drugs," for a time it was popularly believed that humanity might, once and for all, be saved from the scourge of bacterial infections. In fact, during the 1980s and early 1990s, many large pharmaceutical companies cut back or eliminated antibiotics research and development. They believed that infectious disease caused by bacteria finally had been conquered and that markets for new drugs were limited. Unfortunately, this belief was overly optimistic.

The tide is beginning to turn in favor of the bacteria as reports of drug resistant bacteria become more frequent. The United States Centers for Disease Control announced that one of the most powerful known antibiotics, vancomycin, was unable to treat an infection of the common *Staphylococcus aureus* (staph). This organism is commonly found in our environment and is responsible for many nosocomial infections. The import of this announcement becomes clear when one considers that vancomycin was used for years to treat infections caused by stubborn strains of bacteria, like staph. In short, the bacteria are becoming resistant to our most powerful antibiotics. If this trend continues, it is conceivable that we will return to a time when what are presently considered minor bacterial infections are fatal diseases.

There are a number of causes for the predicament in which practitioners of medical arts find themselves. Over-prescription and improper prescription habits by some physicians have caused an indiscriminate increase in the availability of antibiotics to the public. The patient is also partly responsible, for even in instances where an antibiotic is the appropriate treatment, patients will often improperly use the drug, the result being yet another population of bacteria that is resistant, in whole or in part, to traditional antibiotics.

The bacterial scourges that have haunted humanity remain, in spite of the development of modern scientific practices to deal with the diseases that they cause. Drug resistant bacteria are now advancing on the health of humanity. A new generation of antibiotics to once again deal with the pending health threat that bacteria present is required.

Discovery of New Antibiotics

As more and more bacterial strains become resistant to the panel of available antibiotics, new compounds are required. In the past, practitioners of pharmacology would have to rely upon traditional methods of drug discovery to generate novel, safe and efficacious compounds for the treatment of disease. Traditional drug discovery methods involve blindly testing potential drug candidate-molecules, often selected at random, in the hope that one might prove to be an effective treatment for some disease. The process is painstaking and laborious, with no guarantee of success. Today, the average cost to discover and develop a new drug is nearly US $500 million, and the average time is 15 years from laboratory to patient. Improving this process, even incrementally, would represent a huge advance in the generation of novel antimicrobial agents.

Newly emerging practices in drug discovery utilize a number of biochemical techniques to provide for directed approaches to creating new drugs, rather than discovering them at random. For example, gene sequences and proteins encoded thereby that are required for the proliferation of an organism make for excellent targets since exposure of bacteria to compounds active against these targets would result in the inactivation of the organism. Once a target is identified, biochemical analysis of that target can be used to discover or to design molecules that interact with and alter the functions of the target. Using physical and computational techniques, to analyze structural and biochemical targets in order to derive compounds that interact with a target is called rational drug design and offers great future potential. Thus, emerging drug discovery practices use molecular modeling techniques, combinatorial chemistry approaches, and other means to produce and screen and/or design large numbers of candidate compounds.

Nevertheless, while this approach to drug discovery is clearly the way of the future, problems remain. For example, the initial step of identifying molecular targets

for investigation can be an extremely time consuming task. It may also be difficult to design molecules that interact with the target by using computer modeling techniques. Furthermore, in cases where the function of the target is not known or is poorly understood, it may be difficult to design assays to detect molecules that interact with and alter the functions of the target. To improve the rate of novel drug discovery and development, methods of identifying important molecular targets in pathogenic microorganisms and methods for identifying molecules that interact with and alter the functions of such molecular targets are urgently required.

Escherichia coli represents an excellent model system to understand bacterial biochemistry and physiology. The estimated 4288 genes scattered along the 4.6×10^6 base pairs of the *Escherichia coli* (*E. coli*) chromosome offer tremendous promise for the understanding of bacterial biochemical processes. In turn, this knowledge will assist in the development of new tools for the diagnosis and treatment of bacteria-caused human disease. The entire *E. coli* genome has been sequenced, and this body of information holds a tremendous potential for application to the discovery and development of new antibiotic compounds. Yet, in spite of this accomplishment, the general functions or roles of many of these genes are still unknown. For example, the total number of proliferation-required genes contained within the *E. coli* genome is unknown, but has been variously estimated at around 200 to 700 (Armstrong, K.A. and Fan, D.P. Essential Genes in the *metB-malB* Region of *Escherichia coli* K12, 1975, J. Bacteriol. 126: 48-55).

Novel, safe and effective antimicrobial compounds are needed in view of the rapid rise of antibiotic resistant microorganisms. However, prior to this invention, the characterization of even a single bacterial gene was a painstaking process, requiring years of effort. Accordingly, there is an urgent need for more novel methods to identify and characterize bacterial genomic sequences that encode gene products required for proliferation and for methods to identify molecules that interact with and alter the functions of such genes and gene products.

SUMMARY OF THE INVENTION

One embodiment of the present invention is a purified or isolated nucleic acid sequence consisting essentially of one of the sequence of nucleotides of SEQ ID NOs:

1-93, wherein expression of said nucleic acid in a microorganism is capable of inhibiting the proliferation of a microorganism. The nucleic acid sequence may have as sequence of nucleotides complementary to at least a portion of the nucleotide sequence of the coding strand of a gene whose expression is required for proliferation of a microorganism. The nucleic acid may have a nucleotide sequence complementary to at least a portion of the nucleotide sequence of an RNA required for proliferation of a microorganism. The nucleotide sequence of the RNA may encode more than one gene product.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a fragment of one of the nucleotide sequences of SEQ ID NOs.: 1-93, said fragment selected from the group consisting of fragments comprising at least 10, at least 20, at least 25, at least 30, at least 50 and more than 50 consecutive nucleotides of the nucleotide sequence of one of SEQ ID NOs: 1-93.

Another embodiment of the present invention is a vector comprising a promoter operably linked to the nucleic acid sequences of each of the preceding paragraphs. The promoter may be active in a microorganism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

Another embodiment of the present invention is a host cell containing the vectors of the preceding paragraph.

Another embodiment of the present invention is a purified or isolated nucleic acid consisting essentially of the coding sequence of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

Another embodiment of the present invention is a fragment of the nucleic acid of the preceding paragraph, said fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

Another embodiment of the present invention is a vector comprising a promoter operably linked to the nucleic acid of the preceding two paragraphs.

Another embodiment of the present invention is a purified or isolated antisense nucleic acid comprising a nucleic acid sequence complementary to at least a portion of an intragenic sequence, intergenic sequence, sequences spanning at least a portion of two or more genes, 5' noncoding region, or 3' noncoding region within an operon comprising a proliferation-required gene whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a nucleic acid having at least 70% identity to a sequence selected from the group consisting of SEQ ID NOs.: 1-93, fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93, the sequences complementary to SEQ ID NOs.: 1-93 and the sequences complementary to fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93 as determined using BLASTN version 2.0 with the default parameters. The nucleic acid may be from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae,*

Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

Another embodiment of the present invention is a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93. The polypeptide may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a host cell containing the vector of the preceding paragraph.

Another embodiment of the present invention is a purified or isolated polypeptide comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides. The polypeptide may comprise a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a purified or isolated polypeptide comprising a polypeptide having at least 25% identity to a polypeptide whose expression is inhibited by a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid selected

from the group consisting of SEQ ID NOs.: 1-93 as determined using FASTA version 3.0t78 with the default parameters. The polypeptide may have at least 25% identity to a polypeptide comprising one of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 as determined using FASTA version 3.0t78 with the default parameters.

Another embodiment of the present invention is an antibody capable of specifically binding one of the polypeptides of the preceding paragraph.

Another embodiment of the present invention is a method of producing a polypeptide, comprising introducing a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 into a cell and expressing said polypeptide. The method may further comprise the step of isolating said polypeptide. The polypeptide may comprise a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method of inhibiting proliferation of a microorganism comprising inhibiting the activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method for identifying a compound which influences the activity of a gene product required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting said gene product with a candidate compound and determining whether said compound influences the activity of said gene product. The gene product may be a polypeptide and said activity may be an enzymatic

activity. The gene product may be a polypeptide and said activity may be a carbon compound catabolism activity. The gene product may be a polypeptide and said activity may be a biosynthetic activity. The gene product may be a polypeptide and said activity may be a transporter activity. The gene product may be a polypeptide and said activity may be a transcriptional activity. The gene product may be a polypeptide and said activity may be a DNA replication activity. The gene product may be a polypeptide and said activity my be a cell division activity. The gene product may be a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound or nucleic acid having the ability to reduce the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) providing a target that is a gene or RNA, wherein said target comprises a nucleic acid encoding said gene product;

(b) contacting said target with a candidate compound or nucleic acid; and

(c) measuring an activity of said target.

The target may be a messenger RNA molecule and said activity may be translation of said messenger RNA. The target may be a messenger RNA molecule and said activity may be transcription of a gene encoding said messenger RNA. The target may be a gene and said activity may be transcription of said gene. The target may be a nontranslated RNA and said activity may be processing or folding of said nontranslated RNA or assembly of said nontranslated RNA into a protein/RNA complex. The target gene or RNA may encode a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound or nucleic acid identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound which reduces the activity or level of a gene product required for proliferation of a microorganism, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising the steps of:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding said gene product in a cell to reduce the activity or amount of said gene product in said cell, thereby producing a sensitized cell;

(b) contacting said sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

The determining step may comprise determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell. The cell may be selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The cell may be a Gram negative bacterium. The cell may be an *E. coli* cell. The cell may be from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling

within the genera of any of the above species. The antisense nucleic acid may be transcribed from an inducible promoter. The method may further comprise the step of contacting said cell with a concentration of inducer which induces said antisense nucleic acid to a sub-lethal level. Growth inhibition may be measured by monitoring optical density of a culture growth solution. The gene product may be a polypeptide. The polypeptide may comprise a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479. The gene product may be an RNA.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for inhibiting cellular proliferation comprising introducing a compound with activity against a gene whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a compound with activity against the product of said gene into a population of cells expressing said gene. The compound may be an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof. The proliferation inhibiting portion of one of SEQ ID NOs.: 1-93 may be a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 51 consecutive nucleotides of one of SEQ ID NOs.: 1-93. The population may be a population selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The population may be a population of Gram negative bacteria. The population may be a population of *E. coli* cells. The population may be a population selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa,*

Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species. The gene may encode a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a preparation comprising an effective concentration of an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier. The proliferation-inhibiting portion of one of SEQ ID NOs.: 1-93 may comprise at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a method for inhibiting the activity or expression of a gene in an operon required for proliferation wherein the activity or expression of at least one gene in said operon is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting a cell in a cell population with an antisense nucleic acid comprising at least a proliferation-inhibiting portion of said operon. The antisense nucleic acid comprises a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a proliferation inhibiting portion thereof.

The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a plasmid which expresses said antisense nucleic acid into said cell population. The cell may be contacted with said antisense nucleic acid by introducing a phage which expresses said antisense nucleic acid into said cell population. The cell may be contacted with said antisense nucleic acid by expressing said antisense nucleic acid from the chromosome of cells in said cell population. The cell may be contacted with said antisense nucleic acid by introducing a promoter adjacent to a chromosomal copy of said antisense nucleic acid such that said promoter directs the synthesis of said antisense nucleic acid. The cell may be contacted with said antisense nucleic acid by introducing a retron which expresses said antisense nucleic acid into said cell

population. The cell may be contacted with said antisense nucleic acid by introducing a ribozyme into said cell-population, wherein a binding portion of said ribozyme is complementary to said antisense oligonucleotide. The cell may be contacted with said antisense nucleic acid by introducing a liposome comprising said antisense oligonucleotide into said cell. The cell may be contacted with said antisense nucleic acid by electroporation of said antisense nucleic acid. The antisense nucleic acid may be a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93. The antisense nucleic acid may be an oligonucleotide.

Another embodiment of the present invention is a method for identifying a gene which is required for proliferation of a microorganism comprising:

(a) contacting a microorganism other than *E. coli* with a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93;

(b) determining whether said nucleic acid inhibits proliferation of said microorganism; and

(c) identifying the gene in said microorganism which is inhibited by said nucleic acid.

The microorganism may be a Gram negative bacterium. The microorganism may be selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling

within the genera of any of the above species. The method may further comprise introducing said nucleic acid into a vector functional in said microorganism prior to introducing said inhibitory nucleic acid into said microorganism.

Another embodiment of the present invention is a method for identifying a compound having the ability to inhibit proliferation of a microorganism comprising:

(a) identifying in a first microorganism a homolog of a gene or gene product present in a second microorganism which is different than said first microorganism, wherein the activity or level of said gene or gene product is inhibited by a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 ;

(b) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homolog in said first microorganism;

(c) contacting said first microorganism with a sub-lethal level of said inhibitory nucleic acid, thus sensitizing said first microorganism;

(d) contacting the sensitized microorganism of step (c) with a compound; and

(e) determining whether said compound inhibits proliferation of said sensitized microorganism.

The determining step may comprise determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism. Step (a) may comprise identifying a homologous nucleic acid to a gene or gene product whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 or a nucleic acid encoding a homologous polypeptide to a polypeptide whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 by using an algorithm selected from the group consisting of BLASTN version 2.0 with the default parameters and FASTA version 3.0t78 algorithm with the default parameters to identify said homologous nucleic acid or said nucleic acid encoding a homologous polypeptide in a database. Step (a) may comprise identifying a homologous nucleic acid or a nucleic acid encoding a homologous polypeptide by identifying nucleic acids which hybridize to said first gene. Step (a) may comprise expressing a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 in said microorganism. The inhibitory nucleic acid may be an antisense nucleic acid. The

inhibitory nucleic acid may comprise an antisense nucleic acid to a portion of said homolog. The inhibitory nucleic acid may comprise an antisense nucleic acid to a portion of the operon encoding said homolog. The step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid may comprise directly contacting said microorganism with said inhibitory nucleic acid. The step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid may comprise expressing an antisense nucleic acid to said homolog in said microorganism. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the method of the preceding paragraph.

Another embodiment of the present invention is a method of identifying a compound having the ability to inhibit proliferation comprising:

(a) contacting a microorganism other than *E. coli* with a sub-lethal level of a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 or a portion thereof which inhibits the proliferation of *E. coli*, thus sensitizing said microorganism;

(b) contacting the sensitized microorganism of step (a) with a compound; and

(c) determining whether said compound inhibits proliferation of said sensitized microorganism.

The determining step may comprise determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound having activity against a biological pathway required for proliferation comprising:

(a) sensitizing a cell by expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for

proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, in said cell to reduce the activity or amount of said gene product;

(b) contacting the sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

The determining step may comprise determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell. The cell may be selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The cell may be a Gram negative bacterium. The Gram negative bacterium may be *E. coli*. The cell may be selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species. The antisense nucleic acid may be transcribed from an inducible promoter. The method may further comprise contacting the cell with an agent which induces expression of said antisense nucleic acid from said inducible promoter, wherein said antisense nucleic acid is expressed at a sub-lethal level. The inhibition of proliferation may be measured by monitoring the optical density of a liquid culture. The gene product may comprise a polypeptide comprising a

sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound having the ability to inhibit cellular proliferation comprising:

(a) contacting a cell with an agent which reduces the activity or level of a gene product required for proliferation of said cell, wherein said gene product is a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93;

(b) contacting said cell with a compound; and

(c) determining whether said compound reduces proliferation of said contacted cell.

The determining step may comprise determining whether said compound reduces proliferation of said contacted cell to a greater extent than said compound reduces proliferation of cells which have not been contacted with said agent. The agent which reduces the activity or level of a gene product required for proliferation of said cell may comprise an antisense nucleic acid to a gene or operon required for proliferation. The agent which reduces the activity or level of a gene product required for proliferation of said cell may comprise a compound known to inhibit growth or proliferation of a microorganism. The cell may contain a mutation which reduces the activity or level of said gene product required for proliferation of said cell. The mutation may be a temperature sensitive mutation. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the method of the preceding paragraph.

Another embodiment of the present invention is a method for identifying the biological pathway in which a proliferation-required gene or its gene product lies, wherein said gene or gene product comprises a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid which inhibits the activity of said proliferation-required gene or gene product in a cell;

(b) contacting said cell with a compound known to inhibit growth or proliferation of a microorganism, wherein the biological pathway on which said compound acts is known; and

(c) determining whether said cell is sensitive to said compound.

The determining step may comprise determining whether said cell has a substantially greater sensitivity to said compound than a cell which does not express said sub-lethal level of said antisense nucleic acid and wherein said gene or gene product lies in the same pathway on which said compound acts if said cell expressing said sub-lethal level of said antisense nucleic acid has a substantially greater sensitivity to said compound than said cell which does not express said sub-lethal level of said antisense nucleic acid. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method for determining the biological pathway on which a test compound acts comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a proliferation-required nucleic acid in a cell, wherein the activity or expression of said proliferation-required nucleic acid is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and wherein the biological pathway in which said proliferation-required nucleic acid or a protein encoded by said proliferation-required polypeptide lies is known,

(b) contacting said cell with said test compound; and

(c) determining whether said cell is sensitive to said test compound.

The determining step may comprise determining whether said cell has a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said antisense nucleic acid. The method may further comprise:

(d) expressing a sub-lethal level of a second antisense nucleic acid complementary to a second proliferation-required nucleic acid in a second cell, wherein said second proliferation-required nucleic acid is in a different biological pathway than said proliferation-required nucleic acid in step (a); and

(e) determining whether said second cell does not have a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said second antisense nucleic acid, wherein said test compound is specific for the biological pathway against which the antisense nucleic acid of step (a) acts if said second cell does not have substantially greater sensitivity to said test compound.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a compound which interacts with a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

Another embodiment of the present invention is a compound which interacts with a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

Another embodiment of the present invention is a method for manufacturing an antibiotic comprising the steps of screening one or more candidate compounds to identify a compound that reduces the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and manufacturing the compound so identified. The screening step may comprise performing any one of the methods of identifying a compound described above.

Another embodiment of the present invention is a method for inhibiting proliferation of a microorganism in a subject comprising administering a compound that reduces the activity or level of a gene product required for proliferation of said microorganism, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 to said subject. The method of subject may be selected from the group consisting of vertebrates, mammals, avians, and human beings. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an IPTG dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing either an antisense clone to the *E. coli rplW* gene (AS-*rplW*) which encodes a ribosomal protein required for protein synthesis and essential for cell proliferation, or an antisense clone to the *elaD* gene (AS-*elaD*) which is not known to be involved in protein synthesis and which is also essential for proliferation.

Figure 2A is a tetracycline dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing antisense to the *rplW* gene (AS-*rplW*) which was carried out in the presence of 0, 20 or 50 μM IPTG.

Figure 2B is a tetracycline dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing antisense to the *elaD* gene (AS-*elaD*) which was carried out in the presence of 0, 20 or 50 μM IPTG.

Figure 3 is a graph showing the fold increase in tetracycline sensitivity of *E. coli* transfected with antisense clones to essential ribosomal protein genes *L23* (AS-*rplW*) and *L7/L12* and *L10* (AS-*rplLrplJ*). Antisense clones to genes known not to be involved in protein synthesis (*atpB/E*(AS-*atpB/E*), *visC* (AS-*visC*), *elaD* (AS-*elaD*), *yohH* (AS-*yohH*) are much less sensitive to tetracycline.

Definitions

By "biological pathway" is meant any discrete cell function or process that is carried out by a gene product or a subset of gene products. Biological pathways include enzymatic, biochemical and metabolic pathways as well as pathways involved in the production of cellular structures such as cell walls. Biological pathways that are usually required for proliferation of microorganisms include, but are not limited to, cell division, DNA synthesis and replication, RNA synthesis (transcription), protein synthesis (translation), protein processing, protein transport, fatty acid biosynthesis, cell wall synthesis, cell membrane production, synthesis and maintenance, and the like.

By "inhibit activity of a gene or gene product" is meant having the ability to interfere with the function of a gene or gene product in such a way as to decrease expression of the gene or to reduce the level or activity of a product of the gene. Agents which inhibit the activity of a gene include agents that inhibit transcription of the gene, agents that inhibit processing of the transcript of the gene, agents that reduce the

stability of the transcript of the gene, and agents that inhibit translation of the mRNA transcribed from the gene. In microorganisms, agents which inhibit the activity of a gene can act to decrease expression of the operon in which the gene resides or alter the folding or processing of operon RNA so as to reduce the level or activity of the gene product. The gene product can be a non-translated RNA such as ribosomal RNA, a translated RNA (mRNA) or the protein product resulting from translation of the gene mRNA. Of particular utility to the present invention are antisense RNAs that have activities against the operons or genes to which they specifically hybridze.

By "activity against a gene product" is meant having the ability to inhibit the function or to reduce the level or activity of the gene product in a cell.

By "activity against a protein" is meant having the ability to inhibit the function or to reduce the level or activity of the protein in a cell.

By "activity against a nucleic acid" is meant having the ability to inhibit the function or to reduce the level or activity of the nucleic acid in a cell.

By "activity against a gene" is meant having the ability to inhibit the function or expression of the gene in a cell.

By "activity against an operon" is meant having the ability to inhibit the function or reduce the level of one or more products of the operon in a cell.

By "antibiotic" is meant an agent which inhibits the proliferation of a microorganism.

By "*E. coli or Escherichia coli*" is meant *Escherichia coli* or any organism previously categorized as a species of *Shigella* including *Shigella boydii, Shigella flexneri, Shigella dysenteriae, Shigella sonnei, Shigella 2A.*

By "identifying a compound" is meant to screen one or more compounds in a collection of compounds such as a combinatorial chemical library or other library of chemical compounds or to characterize a single compound by testing the compound in a given assay and determining whether it exhibits the desired activity.

By "inducer" is meant an agent or solution which, when placed in contact with a microorganism, increases transcription from a desired promoter.

As used herein, "nucleic acid" means DNA, RNA, or modified nucleic acids. Thus, the terminology "the nucleic acid of SEQ ID NO: X" includes both the DNA sequence of SEQ ID NO: X and an RNA sequence in which the thymidines in the DNA

sequence have been substituted with uridines in the RNA sequence and in which the deoxyribose backbone of the DNA sequence has been substituted with a ribose backbone in the RNA sequence. Modified nucleic acids are nucleic acids having nucleotides or structures which do not occur in nature, such as nucleic acids in which the internucleotide phosphate residues with methylphosphonates, phosphorothioates, phosphoramidates, and phosphate esters. Nonphosphate internucleotide analogs such as siloxane bridges, carbonate brides, thioester bridges, as well as many others known in the art may also be used in modified nucleic acids. Modified nucleic acids may also comprise, α-anomeric nucleotide units and modified nucleotides such as 1,2-dideoxy-d-ribofuranose, 1,2-dideoxy-1-phenylribofuranose, and N^4, N^4-ethano-5-methyl-cytosine are contemplated for use in the present invention. Modified nucleic acids may also be peptide nucleic acids in which the entire deoxyribose-phosphate backbone has been exchanged with a chemically completely different, but structurally homologous, polyamide (peptide) backbone containing 2-aminoethyl glycine units.

As used herein, "sub-lethal" means a concentration of an agent below the concentration required to inhibit all cell growth.

DETAILED DESCRIPTION OF THE INVENTION

The present invention describes a group of *E. coli* genes and gene families required for growth and/or proliferation. A proliferation-required gene or gene family is one where, in the absence of a gene transcript and/or gene product, growth or viability of the microorganism is reduced or eliminated. Thus, as used herein the terminology "proliferation-required" or "required for proliferation" encompasses instances where the absence of a gene transcript and/or gene product completely eliminates cell growth as well as instances where the absence of a gene transcript and/or gene product merely reduces cell growth. These proliferation-required genes can be used as potential targets for the generation of new antimicrobial agents. To achieve that goal, the present invention also encompasses novel assays for analyzing proliferation-required genes and for identifying compounds which interact with the gene products of the proliferation-required genes. In addition, the present invention contemplates the expression of genes and the purification of the proteins encoded by the nucleic acid sequences identified as required proliferation genes and reported herein. The purified proteins can be used to generate reagents and

screen small molecule libraries or other candidate compound libraries for compounds that can be further developed to yield novel antimicrobial compounds. The present invention also describes methods for identification of homologous genes or polypeptides in organisms other than *E. coli*.

The present invention utilizes a novel method to identify proliferation-required *E. coli* sequences. Generally, a library of nucleic acid sequences from a given source are subcloned or otherwise inserted into an inducible expression vector, thus forming an expression library. Although the insert nucleic acids may be derived from the chromosome of the organism into which the expression vector is to be introduced, because the insert is not in its natural chromosomal location, the insert nucleic acid is an exogenous nucleic acid for the purposes of the discussion herein. The term expression is defined as the production of an RNA molecule from a gene, gene fragment, genomic fragment, or operon. Expression can also be used to refer to the process of peptide or polypeptide synthesis. An expression vector is defined as a vehicle by which a ribonucleic acid (RNA) sequence is transcribed from a nucleic acid sequence carried within the expression vehicle. The expression vector can also contain features that permit translation of a protein product from the transcribed RNA message expressed from the exogenous nucleic acid sequence carried by the expression vector. Accordingly, an expression vector can produce an RNA molecule as its sole product or the expression vector can produce a RNA molecule that is ultimately translated into a protein product.

Once generated, the expression library containing the exogenous nucleic acid sequences is introduced into an *E. coli* population to search for genes that are required for bacterial proliferation. Because the library molecules are foreign to the population of *E. coli*, the expression vectors and the nucleic acid segments contained therein are considered exogenous nucleic acid.

Expression of the exogenous nucleic acid fragments in the test population of *E. coli* containing the expression vector library is then activated. Activation of the expression vectors consists of subjecting the cells containing the vectors to conditions that result in the expression of the exogenous nucleic acid sequences carried by the expression vector library. The test population of *E. coli* cells is then assayed to determine the effect of expressing the exogenous nucleic acid fragments on the test population of cells. Those

expression vectors that, upon activation and expression, negatively impact the growth of the *E. coli* screen population are identified, isolated, and purified for further study.

A variety of assays are contemplated to identify nucleic acid sequences that negatively impact growth upon expression. In one embodiment, growth in *E. coli* cultures expressing exogenous nucleic acid sequences is compared to growth in cultures not expressing these sequences. Optical density is used to monitor the extent of growth. Alternatively, enzymatic assays can be used to determine bacterial growth rates to identify exogenous nucleic acid sequences of interest. Colony size, colony morphology, and cell morphology are additional factors used to evaluate growth of the host cells. Those cultures that fail to grow or grow at a reduced rate under expression conditions are identified as containing an expression vector encoding a nucleic acid fragment that negatively affects a proliferation-required gene.

Once exogenous nucleic acid sequences of interest are identified, they are analyzed. The first step of the analysis is to acquire the nucleic acid sequence of the nucleic acid fragment of interest. To achieve this end, the insert in those expression vectors identified as containing a sequence of interest is sequenced, using standard techniques well known in the art. The next step of the process is to determine the source of the nucleic acid sequence.

Determination of sequence source is achieved by comparing the obtained sequence data with known sequences in various genetic databases. The sequences identified are used to probe these gene databases. The result of this procedure is a list of exogenous nucleic acid sequences corresponding to a list that includes novel bacterial genes required for proliferation as well as genes previously identified as required for proliferation.

The number of DNA and protein sequences available in database systems has been growing exponentially for years. For example, at the end of 1998, the complete sequences of *Caenorhabditis elegans*, *Saccharomyces cerevisiae* and nineteen bacterial genomes, including *E. coli* were available. This sequence information is stored in a number of databanks, such as GenBank (the National Center for Biotechnology Information (NCBI), and is publicly available for searching.

A variety of computer programs are available to assist in the analysis of the sequences stored within these databases. FASTA, (W. R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA" Methods in Enzymology

183:63- 98), Sequence Retrieval System (SRS), (Etzold & Argos, SRS an indexing and retrieval tool for flat file data libraries. Comput. Appl. Biosci. 9:49-57, 1993) are two examples of computer programs that can be used to analyze sequences of interest. In one embodiment of the present invention, the BLAST family of computer programs, which includes BLASTN version 2.0 with the default parameters, or BLASTX version 2.0 with the default parameters, is used to analyze nucleic acid sequences.

BLAST, an acronym for "Basic Local Alignment Search Tool," is a family of programs for database similarity searching. The BLAST family of programs includes: BLASTN, a nucleotide sequence database searching program, BLASTX, a protein database searching program where the input is a nucleic acid sequence; and BLASTP, a protein database searching program where the input is an amino acid sequence. BLAST programs embody a fast algorithm for sequence matching, rigorous statistical methods for judging the significance of matches, and various options for tailoring the program for special situations. Assistance in using the program can be obtained by e-mail at blast@ncbi.nlm.nih.gov.

Bacterial genes are often transcribed in polycistronic groups. These groups comprise operons, which are a collection of genes and intergenic sequences. The genes of an operon are co-transcribed and often have related functions. Given the nature of the screening protocol, it is possible that the identified exogenous nucleic acid sequence corresponds to a gene or portion thereof with or without adjacent noncoding sequences, an intragenic sequence (i.e. a sequence within a gene), an intergenic sequence (i.e. a sequence between genes), a sequence spanning at least a portion of two or more genes, a 5' noncoding region or a 3' noncoding region located upstream or downstream from the actual sequence that is required for bacterial proliferation. Accordingly, determining which gene(s) that is encoded within the operons is individually required for proliferation is often desirable.

In one embodiment of the present invention, an operon is dissected to determine which gene or genes are required for proliferation. For example, the RegulonDB DataBase described by Huerta et al. (*Nucl. Acids Res.* **26**:55-59, 1998), which may also be found on the website http://www.cifn.unam.mx/Computational_Biology/regulondb/, may be used to identify the boundaries of operons encoded within microbial genomes. A number of techniques that are well known in the art can then be used to dissect the operon.

In one aspect of this embodiment, gene disruption by homologous recombination is used to individually inactivate the genes of an operon that is thought to contain a gene required for proliferation.

Several gene disruption techniques have been described for the replacement of a functional gene with a mutated, non-functional (null) allele. These techniques generally involve the use of homologous recombination. The method described by Link et al. (J. Bacteriol 1997 179:6228; incorporated herein by reference in it's entirety) serves as an excellent example of these methods as applicable to disruption of genes in *E. coli*. This technique uses crossover PCR to create a null allele with an in-frame deletion of the coding region of a target gene. The null allele is constructed in such a way that sequences adjacent to the wild type gene (ca. 500 bp) are retained. These homologous sequences surrounding the deletion null allele provide targets for homologous recombination so that the wild type gene on the *E. coli* chromosome can be replaced by the constructed null allele.

The crossover PCR amplification product is subcloned into the vector pKO3, the features of which include a chloramphenicol resistance gene, the counter-selectable marker *sacB*, and a temperature sensitive autonomous replication function. Following transformation of an *E. coli* cell population with such a vector, selection for cells that have undergone homologous recombination of the vector into the chromosome is achieved by growth on chloramphenicol at the non-permissive temperature of 43°C. Under these conditions, autonomous replication of the plasmid cannot occur and cells are resistant to chloramphinicol only if the chloramphenicol resistance gene has been integrated into the chromosome. Usually a single crossover event is responsible for this integration event such that the *E. coli* chromosome now contains a tandem duplication of the target gene consisting of one wild type allele and one deletion null allele separated by vector sequence.

This new *E. coli* strain containing the tandem duplication can be maintained at permissive temperatures in the presence of drug selection (chloramphenicol). Subsequently, cells of this new strain are cultured at the permissive temperature 30°C without drug selection. Under these conditions, the chromosome of some of the cells within the population will have undergone an internal homologous recombination event resulting in removal of the plasmid sequences. Subsequent culturing of the strain in

growth medium lacking chloramphenicol but containing sucrose is used to select for such recombinative resolutions. In the presence of the counter-selectable marker *sacB*, sucrose is rendered into a toxic metabolite. Thus, cells that survive this counter-selection have lost both the plasmid sequences from the chromosome and the autonomously replicating plasmid that results as a byproduct of recombinative resolution.

There are two possible outcomes of the above recombinative resolution via homologous recombination. Either the wild type copy of the targeted gene is retained on the chromosome or the mutated null allele is retained on the chromosome. In the case of an essential gene, a single copy of the null allele would be lethal and such cells should not be obtained by the above procedure when applied to essential genes. In the case of a non-essential gene, roughly equal numbers of cells containing null alleles and cells containing wild type alleles should be obtained. Thus, the method serves as a test for essentiality of the targeted gene: when applied to essential genes, only cells with a wild type allele on the chromosome will be obtained.

Other techniques have also been described for the creation of disruption mutations in *E. coli*. For example, Link et al. also describe inserting an in-frame sequence tag concommitantly with an in-frame deletion in order to simplify analysis of recombinants obtained. Further, Link et al. describe disruption of genes with a drug resistance marker such as a kanamycin resistance gene. Arigoni et al., (Arigoni, F. et al. A Genome-based Approach for the Identification of Essential Bacterial Genes, Nature Biotechnology 16: 851-856, the disclosure of which is incorporated herein by reference in its entirety) describe the use of gene disruption combined with engineering a second copy of a test gene such that the expression of the gene is regulated by and inducible promoter such as the arabinose promoter to test the essentiality of the gene. Many of these techniques result in the insertion of large fragments of DNA into the gene of interest, such as a drug selection marker. An advantage of the technique described by Link et al. is that it does not rely on an insertion into the gene to cause a functional defect, but rather results in the precise removal of the coding region. This insures the lack of polar effects on the expression of genes downstream from the target gene.

Recombinant DNA techniques can be used to express the entire coding sequences of the gene identified as required for proliferation, or portions thereof. The over-expressed

proteins can be used as reagents for further study. The identified exogenous sequences are isolated, purified, and cloned into a suitable expression vector using methods well known in the art. If desired, the nucleic acids can contain the sequences encoding a signal peptide to facilitate secretion of the expressed protein.

Expression of fragments of the bacterial genes identified as required for proliferation is also contemplated by the present invention. The fragments of the identified genes can encode a polypeptide comprising at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 75, or more than 75 consecutive amino acids of a gene complementary to one of the identified sequences of the present invention. The nucleic acids inserted into the expression vectors can also contain sequences upstream and downstream of the coding sequence.

When expressing the coding sequence of an entire gene identified as required for bacterial proliferation or a fragment thereof, the nucleic acid sequence to be expressed is operably linked to a promoter in an expression vector using conventional cloning technology. The expression vector can be any of the bacterial, insect, yeast, or mammalian expression systems known in the art. Commercially available vectors and expression systems are available from a variety of suppliers including Genetics Institute (Cambridge, MA), Stratagene (La Jolla, California), Promega (Madison, Wisconsin), and Invitrogen (San Diego, California). If desired, to enhance expression and facilitate proper protein folding, the codon usage and codon bias of the sequence can be optimized for the particular expression organism in which the expression vector is introduced, as explained by Hatfield, et al., U.S. Patent No. 5,082,767, incorporated herein by this reference. Fusion protein expression systems are also contemplated by the present invention.

Following expression of the protein encoded by the identified exogenous nucleic acid sequence, the protein is purified. Protein purification techniques are well known in the art. Proteins encoded and expressed from identified exogenous nucleic acid sequences can be partially purified using precipitation techniques, such as precipitation with polyethylene glycol. Alternatively, epitope tagging of the protein can be used to allow simple one step purification of the protein. Chromatographic methods usable with the present invention can include ion-exchange chromatography, gel filtration, use of hydroxyapaptite columns, immobilized reactive dyes, chromatofocusing, and use of high-

performance liquid chromatography. Electrophoretic methods such one-dimensional gel electrophoresis, high-resolution two-dimensional polyacrylamide electrophoresis, isoelectric focusing, and others are contemplated as purification methods. Also, affinity chromatographic methods, comprising antibody columns, ligand presenting columns and other affinity chromatographic matrices are contemplated as purification methods in the present invention.

The purified proteins produced from the gene coding sequences identified as required for proliferation can be used in a variety of protocols to generate useful antimicrobial reagents. In one embodiment of the present invention, antibodies are generated against the proteins expressed from the identified exogenous nucleic acid sequences. Both monoclonal and polyclonal antibodies can be generated against the expressed proteins. Methods for generating monoclonal and polyclonal antibodies are well known in the art. Also, antibody fragment preparations prepared from the produced antibodies discussed above are contemplated.

In addition, the purified protein, fragments therof, or derivatives thereof may be administered to an individual in a pharmaceutically acceptable carrier to induce an immune response against the protein. Preferably, the immune response is a protective immune response which protects the individual. Methods for determining appropriate dosages of the protein and pharmaceutically acceptable carriers are familiar to those skilled in the art.

Another application for the purified proteins of the present invention is to screen small molecule libraries for candidate compounds active against the various target proteins of the present invention. Advances in the field of combinatorial chemistry provide methods, well known in the art, to produce large numbers of candidate compounds that can have a binding, or otherwise inhibitory effect on a target protein. Accordingly, the screening of small molecule libraries for compounds with binding affinity or inhibitory activity for a target protein produced from an identified gene sequence is contemplated by the present invention.

The present invention further contemplates utility against a variety of other pathogenic organisms in addition to *E. coli*. For example, the invention has utility in identifying genes required for proliferation in prokaryotes and eukaryotes. For example, the invention has utility with protists, such as *Plasmodium* spp .and as *Entamoeba* spp.;

plants; animals, such *and Contracaecum* spp*;* and fungi including *Candida* spp., ((e.g., *Candida albicansCandida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis,*)), *Saccharomyces cerevisiae*, *Cryptococcus neoformans*, and *Aspergillus fumigatus*. In one embodiment of the present invention, monera, specifically bacteria are probed in search of novel gene sequences required for proliferation. This embodiment is particularly important given the rise of drug resistant bacteria.

The numbers of bacterial species that are becoming resistant to existing antibiotics are growing. A partial list of these organisms includes: *Staphylococcus* spp., such as *S. aureus*; *Enterococcus* spp., such as *E. faecalis*; *Pseudomonas* spp., *such as P. aeruginosa*, *Clostridium* spp., such as *C. botulinum or C. difficile, Haemophilus* spp., such as *H. influenzae*, *Enterobacter* spp., such as *E. cloacae, Vibrio* spp., such as *V. cholera; Moraxala* spp., such as *M. catarrhalis*; *Streptococcus* spp., such as *S. pneumoniae, Neisseria* spp., such as *N. gonorrhoeae; Mycoplasma* spp., such as *Mycoplasma pneumoniae; Salmonella typhimurium*; *Helicobacter pylori*; *Escherichia coli;* and *Mycobacterium tuberculosis*. The sequences identified as required for proliferation in the present invention can be used to probe these and other organisms to identify homologous required proliferation genes contained therein.

In one embodiment of the present invention, the nucleic acid sequences disclosed herein are used to screen genomic libraries generated from bacterial species of interest other than *E. coli*. For example, the genomic library may be from *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria*

monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis or any species falling within the genera of any of the above species. Standard molecular biology techniques are used to generate genomic libraries from various microorganisms. In one aspect, the libraries are generated and bound to nitrocellulose paper. The identified exogenous nucleic acid sequences of the present invention can then be used as probes to screen the libraries for homologous sequences. The homologous sequences identified can then be used as targets for the identification of new, antimicrobial compounds with activity against more than one organism.

For example, the preceding methods may be used to isolate nucleic acids having a sequence with at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% identity to a nucleic acid sequence selected from the group consisting of one of the sequences of SEQ ID NOS. 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286, fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof, and the sequences complementary thereto. Identity may be measured using BLASTN version 2.0 with the default parameters. (Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucleic Acid Res. 25: 3389-3402 (1997), the disclosure of which is incorporated herein by reference in its entirety). For example, the homologous polynucleotides may have a coding sequence which is a naturally occurring allelic variant of one of the coding sequences described herein. Such allelic variants may have a substitution, deletion or addition of one or more nucleotides when compared to the nucleic acids of SEQ ID NOs: 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286 or the sequences complementary thereto.

Additionally, the above procedures may be used to isolate nucleic acids which encode polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, or at least 40% identity or similarity to a polypeptide having the sequence of one of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or to a polypeptide whose expression is inhibited by a nucleic acid of one of SEQ ID NOs.: 1-93, or fragments comprising at

least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of the preceding polypeptides as determined using the FASTA version 3.0t78 algorithm with the default parameters. Alternatively, protein identity or similarity may be identified using BLASTP with the default parameters, BLASTX with the default parameters, or TBLASTN with the default parameters. (Alschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucleic Acid Res. 25: 3389-3402 (1997), the disclosure of which is incorporated herein by reference in its entirety).

Alternatively, homologous nucleic acids or polypeptides may be identified by searching a database to identify sequences having a desired level of homology to a nucleic acid or a polypeptide involved in proliferation or an antisense nucleic acid to a nucleic acid involved in microbial proliferation. A variety of such databases are available to those skilled in the art, including GenBank and GenSeq. In some embodiments, the databases are screened to identify nucleic acids or polypeptides having at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, or at least 50%, at least 40% identity or similarity to a nucleic acid or polypeptide involved in proliferation or an antisense nucleic acid involved in proliferation. For example, the database may be screened to identify nucleic acids homologous to one of SEQ ID Nos. 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286, homologous to fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof, or homologous to the sequences complementary to any of the preceding nucleic acids. In other embodiments, the databases are screened to indetify polypeptides having at least 99%, 95%, 90%, 855, 80%, 70%, 60%, 50%, 40%, or at least 25% identity or similarity ot a polypeptide involved in proliferation or a portion thereof. For example, the database may be screened to identify polypeptides homologous to a polypeptide comprising one of SEQ ID NOs. 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479, a polypeptide whose expression is inhibited by a nucleic acid of one of SEQ ID NOs: 1-93, or homologous to fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of any of the preceding polypeptides. In some embodiments, the database may be screened to identify homologous nucleic acids or polypeptides from organisms other than *E. coli*, including

organisms such as *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis,Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* or any species falling within the genera of any of the above species.

In another embodiment, gene expression arrays and microarrays can be employed. Gene expression arrays are high density arrays of DNA samples deposited at specific locations on a glass chip, nylon membrane, or the like. Such arrays can be used by researchers to quantify relative gene expression under different conditions. Gene expression arrays are used by researchers to help identify optimal drug targets, profile new compounds, and determine disease pathways. An example of this technology is found in U.S. Patent No. 5807522, which is hereby incorporated by reference.

It is possible to study the expression of all genes in the genome of a particular microbial organism using a single array. For example, the arrays from Genosys consist of 12 x 24 cm nylon filters containing PCR products corresponding to 4290 ORFs from *E. coli*. 10 ngs of each are spotted every 1.5 mm on the filter. Single stranded labeled cDNAs are prepared for hybridization to the array (no second strand synthesis or

amplification step is done) and placed in contact with the filter. Thus the labeled cDNAs are of "antisense" orientation. Quantitative analysis is done by phosphorimager.

Hybridization of cDNA made from a sample of total cell mRNA to such an array followed by detection of binding by one or more of various techniques known to those in the art results in a signal at each location on the array to which cDNA hybridized. The intensity of the hybridization signal obtained at each location in the array thus reflects the amount of mRNA for that specific gene that was present in the sample. Comparing the results obtained for mRNA isolated from cells grown under different conditions thus allows for a comparison of the relative amount of expression of each individual gene during growth under the different conditions.

Gene expression arrays may be used to analyze the total mRNA expression pattern at various time points after induction of an antisense nucleic acid complementary to a proliferation-required gene. Analysis of the expression pattern indicated by hybridization to the array provides information on whether or not the target gene of the antisense nucleic acid is being affected by antisense induction, how quickly the antisense is affecting the target gene, and for later timepoints, what other genes are affected by antisense expression. For example, if the antisense is directed against a gene for ribosomal protein L7/L12 in the 50S subunit, its targeted mRNA may disappear first and then other mRNAs may be observed to increase, decrease or stay the same. Similarly, if the antisense is directed against a different 50S subunit ribosomal protein mRNA (e.g. L25), that mRNA may disappear first followed by changes in mRNA expression that are similar to those seen with the L7/L12 antisense expression. Thus, the mRNA expression pattern observed with an antinsense nucleic acid complementary to a proliferation required gene may identify other proliferation-required nucleic acids in the same pathway as the target of the antisense nucleic acid. In addition, the mRNA expression patterns observed with candidate drug compounds may be compared to those observed with antisense nucleic acids against a proliferation-required nucleic acid. If the mRNA expression pattern observed with the candidate drug compound is similar to that observed with the antisense nucleic acid, the drug compound may be a promising therapeutic candidate. Thus, the assay would be useful

in assisting in the selection of candidate drug compounds for use in screening methods such as those described below.

In cases where the source of nucleic acid deposited on the array and the source of the nucleic acid being hybridized to the array are from two different organisms, gene expression arrays can identify homologous genes in the two organisms.

The present invention also contemplates additional methods for screening other microorganisms for proliferation-required genes. In this embodiment, the conserved portions of sequences identified as proliferation-required can be used to generate degenerate primers for use in the polymerase chain reaction (PCR). The PCR technique is well known in the art. The successful production of a PCR product using degenerate probes generated from the sequences identified herein would indicate the presence of a homologous gene sequence in the species being screened. This homologous gene is then isolated, expressed, and used as a target for candidate antibiotic compounds. In another aspect of this embodiment, the homologous gene is expressed in an autologous organism or in a heterologous organism in such a way as to alter the level or activity of a homologous gene required for proliferation in the autologous or heterologus organism. In still another aspect of this embodiment, the homologous gene or portion is expressed in an antisense orientation in such a way as to alter the level or activity of a nucleic acid required for proliferation of an autologous or heterologous organism.

The homologous sequences to proliferation-required genes identified using the techniques described herein may be used to identify proliferation-required genes of organisms other than *E. coli*, to inhibit the proliferation of organisms other than *E. coli* by inhibiting the activity or reducing the amount of the identified homologous nucleic acid or polypeptide in the organism other than *E. coli,* or to identify compounds which inhibit the growth of organisms other than *E. coli* as described below.

In another embodiment of the present invention, *E. coli* sequences identified as required for proliferation are transferred to expression vectors capable of function within non-*E coli* species. As would be appreciated by one of ordinary skill in the art, expression vectors must contain certain elements that are species specific. These elements can include promoter sequences, operator sequences, repressor genes, origins of replication,selectable marker genes, ribosomal binding sequences, termination sequences, and others. To use the identified exogenous sequences of the present

invention, one of ordinary skill in the art would know to use standard molecular biology techniques to isolate vectors containing the sequences of interest from cultured bacterial cells, isolate and purify those sequences, and subclone those sequences into an expression vector adapted for use in the species of bacteria to be screened.

Expression vectors for a variety of other species are known in the art. For example, Cao et al. report the expression of steroid receptor fragments in *Staphylococcus aureus*. J. Steroid Biochem Mol Biol. 44(1):1-11 (1993). Also, Pla et al. have reported an expression vector that is functional in a number of relevant hosts including: *Salmonella typhimurium, Pseudomonas putida, and Pseudomonas aeruginosa*. J. Bacteriol. 172(8):4448-55 (1990). These examples demonstrate the existence of molecular biology techniques capable of constructing expression vectors for the species of bacteria of interest to the present invention.

Following the subcloning of the identified nucleic acid sequences into an expression vector functional in the microorganism of interest, the identified nucleic acid sequences are conditionally transcribed to assay for bacterial growth inhibition. Those expression vectors found to contain sequences that, when transcribed, inhibit bacterial growth are compared to the known genomic sequence of the pathogenic microorganism being screened or, if the homologous sequence from the organism being screened is not known, it may be identified and isolated by hybridization to the proliferation-required *E. coli* sequence interest or by amplification using primers based on the proliferation-required *E. coli* sequence of interest as described above.

The antisense sequences from the second organism which are identified as described above may then be operably linked to a promoter, such as an inducible promoter, and introduced into the second organism. The techniques described herein for identifying *E. coli* genes required for proliferation may thus be employed to determine whether the identified sequences from a second organism inhibit the proliferation of the second organism.

Antisense nucleic acids required for the proliferation of organisms other than *E. coli* or the genes corresponding thereto, may also be hybridized to a microarray containing the *E. coli* ORFs to gauge the homology between the *E. coli* sequences and the proliferation-required nucleic acids from other organisms. For example, the proliferation-required nucleic acid may be from *Aspergillus fumigatus, Bacillus*

anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis or any species falling within the genera of any of the above species. The proliferation-required nucleic acids from an organism other than *E. coli* may be hybridized to the array under a variety of conditions which permit hybridization to occur when the probe has different levels of homology to the sequence on the microarray. This would provide an indication of homology across the organisms as well as clues to other possible essential genes in these organisms.

In still another embodiment, the exogenous nucleic acid sequences of the present invention that inhibit bacterial growth or proliferation can be used as antisense therapeutics for killing bacteria. The antisense sequences can be complementary to the proliferation-required genes whose sequence corresponds to the exogenous nucleic acid probes identified here (i.e. the antisense nucleic acid may hybridize to the gene or a portion thereof). Alternatively, antisense therapeutics can be complementary to operons in which proliferation-required genes reside (i.e. the antisense nucleic acid may hybridize to any gene in the operon in which the proliferation-required genes reside). Further, antisense therapeutics can be complementary to a proliferation-required gene or portion thereof with or without adjacent noncoding sequences, an intragenic sequence (i.e. a sequence within a gene), an intergenic sequence (i.e. a sequence between genes), a sequence spanning at least a portion of two or more genes, a 5' noncoding region or a 3' noncoding region located

upstream or downstream from the actual sequence that is required for bacterial proliferation or an operon containing a proliferation-required gene.

In addition to therapeutic applications, the present invention encompasses the use of nucleic acid sequences complementary to sequences required for proliferation as diagnostic tools. For example, nucleic acid probes complementary to proliferation-required sequences that are specific for particular species of microorganisms can be used as probes to identify particular microorganism species in clinical specimens. This utility provides a rapid and dependable method by which to identify the causative agent or agents of a bacterial infection. This utility would provide clinicians the ability to prescribe species specific antimicrobial compounds to treat such infections. In an extension of this utility, antibodies generated against proteins translated from mRNA transcribed from proliferation-required sequences can also be used to screen for specific microorganisms that produce such proteins in a species-specific manner.

The following examples teach the genes of the present invention and a subset of uses for the *E. coli* genes identified as required for proliferation. These examples are illustrative only and are not intended to limit the scope of the present invention.

EXAMPLES

The following examples are directed to the identification and exploitation of *E. coli* genes required for proliferation. Methods of gene identification are discussed as well as a variety of methods to utilize the identified sequences.

Genes Identified as Required for Proliferation of *E. coli*

Exogenous nucleic acid sequences were cloned into an inducible expression vector and assayed for growth inhibition activity. Example 1 describes the examination of a library of exogenous nucleic acid sequences cloned into the IPTG-inducible expression vector pLEX5BA (Krause et al., J. Mol. Biol. 274: 365 (1997), the disclosure of which is incorporated herein by reference in its entirety) or a modified version of pLEX5BA, pLEX5BA-3' in which a synthetic linker containing a T7 terminator was ligated between the PstI and HindIII sites of pLEX5BA. In particular, to construct pLEX5BA-3', the following oligonucleotides were annealed and inserted into the PstI and HindIII sites of pLEX5BA:

5'-GTCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCCTTGAGGGGTTTTTTGA-3' (SEQ ID NO: 480)

5'-AGCTTCAAAAAACCCCTCAAGGACCCGTTTAGAGGCCCCAAGGGGTTAT GCTAGACTGCA-3' (SEQ ID NO: 481)

Random fragments of *E. coli* genomic DNA were generated by DNAseI digestion or sonication, filled in with T4 polymerase, and cloned into the SmaI site of pLEX5BA or pLEX5BA-3'. Upon activation or induction, the expression vectors produced an RNA molecule corresponding to the subcloned exogenous nucleic acid sequences. The RNA product was in an antisense orientation with respect to the *E. coli* genes from which it was originally derived. This antisense RNA then interacted with sense mRNA produced from various *E. coli* genes and interfered with or inhibited the translation of the sense messenger RNA (mRNA) thus preventing protein production from these sense mRNA molecules. In cases where the sense mRNA encoded a protein required for the proliferation, bacterial cells containing an activated expression vector failed to grow or grew at a substantially reduced rate. Similar results have also been obtained in cases where the gene encodes a non-translated RNA, such as a ribosomal RNA.

It will be appreciated that vectors other than pLEX5BA or pLEX5BA-3' may be used to transcribe the genomic DNA inserts. In addition, it will be appreciated that, if desired, pLEX5BA or pLEX5BA-3' may be modified to introduce features such as stop codons in all three reading frames downstream of the genomic DNA inserts to ensure that if the genomic DNA insert encodes a polypeptide (i.e. the insert is in the sense orientation rather than the antisense orientation or the insert is in the antisense orientation but contains a cryptic ORF) translation of the polypeptide will terminate shortly after the genomic insert.

EXAMPLE 1

Inhibition of Bacterial Proliferation after IPTG induction

To study the effects of transcriptional induction in liquid medium, growth curves were carried out by back diluting cultures 1:200 into fresh media with or without 1 mM IPTG and measuring the OD_{450} every 30 minutes (min). To study the effects of transcriptional induction on solid medium, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7 and 10^8 fold dilutions of overnight cultures were prepared. Aliquots of from 0.5 to 3 µl of these dilutions were spotted on selective agar plates with or without 1 mM IPTG. After overnight incubation, the plates were compared to assess the sensitivity of the clones to IPTG.

Of the numerous clones tested, some clones were identified as containing a sequence that inhibited *E. coli* growth after IPTG induction. Accordingly, the gene to which the inserted nucleic acid sequence corresponds, or a gene within the operon containing the inserted nucleic acid, may be required for proliferation in *E. coli*.

Characterization of Isolated Clones Negatively Affecting *E. coli* Proliferation

Following the identification of those inserts that, upon expression, negatively impacted *E. coli* growth or proliferation, the inserts were isolated and subjected to nucleic acid sequence determination.

EXAMPLE 2

Nucleic Acid Sequence Determination of Identified Clones Expressing Nucleic Acid Fragments with Detrimental Effects of *E. coli* Proliferation

The nucleotide sequences for the exogenous identified sequences were determined using plasmid DNA isolated using QIAPREP (Qiagen, Valencia, CA) and methods supplied by the manufacturer. The primers used for sequencing the inserts were 5' - TGTTTATCAGACCGCTT - 3' (SEQ ID NO: 1) and 5' - ACAATTTCACACAGCCTC - 3' (SEQ ID NO: 2). These sequences flank the polylinker in pLEX5BA. Sequence identification numbers (SEQ ID NOs) for the identified inserts are listed in Table I and discussed below.

EXAMPLE 3

Comparison Of Isolated Sequences to Known Sequences

The nucleic acid sequences of the subcloned fragments obtained from the expression vectors discussed above were compared to known *E. coli* sequences in GenBank using BLAST version 1.4 or version 2.0.6 using the following default parameters: Filtering off, cost to open a gap=5, cost to extend a gap=2, penalty for a mismatch in the blast portion of run=-3, reward for a match in the blast portion of run=1, expectation value (e)=10.0, word size=11, number of one-line descriptions=100, number of alignments to show (B)=100. BLAST is described in Altschul, J Mol Biol. 215:403-10 (1990), the disclosure of which is incorporated herein by reference in its entirety. Expression vectors were found to contain nucleic acid sequences in both the sense and antisense orientations. The presence of known genes, open reading frames, and ribosome binding sites was determined by comparison to public databases holding genetic

information and various computer programs such as the Genetics Computer Group programs FRAMES and CODONPREFERENCE. Clones were designated as "antisense" if the cloned fragment was oriented to the promoter such that the RNA transcript produced was complementary to the expressed mRNA (or non-translated RNA) from a chromosomal locus. Clones were designated as "sense" if they coded for an RNA fragment that was identical to a portion of a wild type mRNA from a chromosomal locus.

The sequences described in Examples 1-2 that inhibited bacterial proliferation and contained gene fragments in an antisense orientation are listed in Table I. This table lists each identified sequence by: a sequence identification number; a Molecule Number; a gene to which the identified sequence corresponds, listed according to the National Center for Biotechnology Information (NCBI), Blattner (Science 277:1453-1474(1997); also contains the *E. coli* K-12 genome sequence), or Rudd (Micro. and Mol. Rev. 62:985-1019 (1998)), (both papers are hereby incorporated by reference) nomenclatures. The CONTIG numbers for each identified sequence is shown, as well as the location of the first and last base pairs located on the *E. coli* chromosome. A Molecule Number with a "**" indicates a clone corresponding to an intergenic sequence.

TABLE I

Identified Clones with Corresponding Genes and Operons

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
1	EcXA118a	E1M10000131C06	AE000408	5299	5440	b3310	rplN
2	EcXA118b	E1M10000152F04	AE000408	5300	5452	b3310	rplN
3	EcXA118c	E1M10000152H04	AE000408	5311	5450	b3310	rplN
4	EcXA118d	E1M10000153H03	AE000408	5299	5475	b3310	rplN
5	EcXA119	E1M10000129F10	AE000372	2407	2153	b2883	b2883
6	EcXA120	E1M10000129G04	AE000248	2494	2888	b1509	b1509
7	EcXA121	E1M10000119D03	AE000491	1722	2001	b4191	yjfQ
8	EcXA122a	1029-M5	AE000300	232	1	b2108	yehA
			AE000299	10036	9848	b2107	b2107
9	EcXA122b	X3S208-17	AE000300	186	1	b2108	yehA
			AE000299	10036	9870	b2107	b2107
10	EcXA122c	E1M10000124C07	AE000299	10036	9896	b2107	b2107
			AE000300	78	1	b2108	yehA
11	EcXA122d	E1M10000125D06	AE000300	203	1	b2108	yehA
			AE000299	10036	9892	b2107	b2107
12	EcXA123	E1M10000106A08	AE000410	7832	7968	b3343	yheL
13	EcXA124	E1M10000106D02	AE000408	5951	6225	b3311	rpsQ
			AE000408	5951	6225	b3312	rpmC

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
14	EcXA125	E1M10000124D09	AE000465	7071	7309	b3901	yiiL
15	EcXA126	E1M10000124E02	AE000441	9429	9316	b3644	yicC
16	EcXA127a	E1M10000124E06	AE000369	1145	856	b2851	b2851
17	EcXA127b	E1M10000159A11	AE000369	1145	763	b2851	b2851
18	EcXA128	E1M10000124G04	AE000171	6676	6427	b0679	nagE
19	EcXA129	E1M10000127C09	AE000123	1270	1506	b0135, b0136	yadC, yadK
20	EcXA130	E1M10000146E05	AE000450	9592	9828	b3737, b3738	atpE, atpB
21	EcXA131	E1M10000162C01	AE000383	11211	11350	b3016	b3016
22	EcXA132	E1M10000163A04	AE000243	5440	5650	b1465	narV
23	EcXA133a	E1M10000120F06	AE000408	5273	5425	b3309, b3310	rplX, rplN
24	EcXA133b	E1M10000132H04	AE000408	5263	5422	b3309, b3310	rplX, rplN
25	EcXA133c	E1M10000172C05	AE000408	5190	5433	b3309, b3310	rplX, rplN
26	EcXA134	E1M10000162H06	AE000402	4689	4795	b3229	sspA
27	EcXA136a	E1M10000147B03	AE000373	1	251	b2892	recJ
			AE000372	11967	12144	b2891	prfB
28	EcXA136b	E1M10000155F04	AE000373	1	372	b2892	recJ
			AE000372	12002	12144	b2891	prfB
29	EcXA137	E1M10000142F12	AE000132	9172	8451	b0243, b0244	proA, thrW
30	EcXA138	E1M10000148B09	AE000393	10354	10051	b3121	yhaC
31	EcXA139	E1M10000143A12	AE000189	3600	4024	b0874	b0874
32	EcXA139b	E1M10000169H02	AE000189	3646	4026	b0874	b0874
33	EcXA140a	E1M10000143C10	AE000445	1	260	b3672	b3672
			AE000444	12573	12697	b3671	ilvB
34	EcXA140b	E1M10000151F10	AE000445	1	378	b3672	b3672
			AE000444	12576	12697	b3671	ilvB
35	EcXA141	E1M10000143G09	AE000294	9227	9498	b2035	rfc
36	EcXA142	E1M10000159C10	AE000408	1	199	b3299	rpmJ
			AE000407	10488	10601	b3298	rpsM
37	EcXA143	E1M10000159F03	AE000368	7705	7317	b2846	b2846
38	EcXA144a	E1M10000160A03	AE000257	545	368	b1613	manA
39	EcXA144b	E1M10000132F02	AE000257	585	395	b1613	manA
40	EcXA144c	E1M10000153E04	AE000257	552	395	b1613	manA
41	EcXA145	E1M10000160A06	AE000366	5350	5533	b2827	thyA
42	EcXA146	E1M10000160C02	AE000449	9694	9815	b3725	pstB
43	EcXA147	E1M10000160C03	AE000211	6964	7126	b1111	b1111
44	EcXA148	E1M10000160E10	AE000455	10737	10577	b3794	rffM
45	EcXA149a	E1M10000160F11	AE000374	7930	7680	b2912	ygfA
46	EcXA149b	E1M10000153F05	AE000374	7910	7625	b2912	ygfA
47	EcXA150	E1M10000160H03	AE000442	5556	5377	b3650	spoT
48	EcXA151	E1M10000144E03	AE000429	8099	8352	b3528	dctA
49	EcXA152	E1M10000144G07	AE000308	1086	1385	b2182	bcr
50	EcXA153	E1M10000144B01	AE000477	4261	4055	b4038	yjbI
51	EcXA154	E1M10000150E02	AE000290	5618	5406	b1983	b1983
52	EcXA155	E1M10000150E06	AE000339	116	496	b2522	sseB
53	EcXA156	E1M10000156B08	AE000463	2852	2699	b3874	yihN
54	EcXA157	E1M10000156D07	AE000408	7274	7464	b3314, b3315	rpsC, rplV
55	EcXA158	E1M10000156G12	AE000338	5181	5370	b2519	b2519
56	EcXA159	E1M10000151G10	AE000154	7217	7618	b0482	b0482
57	EcXA160	E1M10000166G06	AE000269	4488	5074	b1744	b1744
58	EcXA161	869.A23	AE000146	3702	3950	b0398	sbcD
59	EcXA162	E1M10000112F05	AE000497	7921	8156	b4267	yjgV
60	EcXA163	E1M10000118B05	AE000447	8832	8741	b3705	yidC

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
61	EcXA164	E1M10000118C04	AE000314	6749	6688	b2243	glpC
62	EcXA165	E1M10000118C05	AE000170	219	290	b0655	ybeJ
63	EcXA166	E1M10000118G06	AE000461	5835	5770	b3859	yihE
64	EcXA167	E1M10000119A05	AE000298	9388	9521	b2093	gatB
65	EcXA168	E1M10000123E09	AE000267	4622	4413	b1721	b1721
66	EcXA169	E1M10000123F11	AE000408	2650	2767	b3303	rpsE
67	EcXA170	E1M10000129G01	AE000143	3523	3913	b0363	b0363
68	EcXA171	E1M10000132G08	AE000189	6799	6699	b0876	b0876
69	EcXA172	E1M10000139B11	AE000323	3697	3531	b2345	b2345
70	EcXA173	E1M10000153A09	AE000424	6886	6968	b3485	yhhJ
71	EcXA174	E1M10000153C04	AE000377	8932	8832	b2945	endA
72	EcXA175	E1M10000154F03	AE000279	7957	7837	b1854	pykA
73	EcXA176	E1M10000157B02	AE000485	4932	5094	b4125	yjdH
74	EcXA177	E1M10000160H05	AE000326	3140	3035	b2378	b2378
75	EcXA178	E1M10000163H01	AE000113	1	266	b0023	rpsT
76	EcXA179	E1M10000164G04	AE000262	233	1	b1665	valV
77	EcXA180	E1M10000168F02	AE000480	3110	3350	b4066	yjcF
78	EcXA181	E1M10000169D08	AE000469	10535	10604	b3956	ppc
79	EcXA182	E1M10000169H07	AE000376	10447	10586	b2935	tktA
80	EcXA183	E1M10000169H08	AE000118	8416	8228	b0084	ftsI
81	EcXA184	E1M10000172C01	AE000358	8498	8844	b2750	cysC
82	EcXA185	E1M10000176C01	AE000407	10053	10119	b3297	rpsK
83	EcXA186	E1M10000176F01	AE000135	987	812	b0270	yagG
84	EcXA187	P319-3.M16	AE000317	2279	1872	b2271	b2271
85	EcXA188	P347.2	AE000211	7845	8005	b1113	b1113
86	EcXA189a	E1M10000168A04	AE000277	2048	1861	b1828	b1828
87	EcXA189b	E1M10000125F09	AE000277	2042	1868	b1828	b1828
88	EcXA189c	E1M10000113F02	AE000277	2050	1857	b1828	b1828
89	EcXA190a	E1M10000161D04	AE000387	49	312	b3052	b3052
90	EcXA190b	E1M10000155B11	AE000387	49	285	b3052	b3052
91	EcXA191a	E1M10000167H09	AE000451	18	220	b3740	gidB
92	EcXA191b	E1M10000120H05	AE000451	1	188	b3740	gidB
93	EcXA192	E1M10000154H07	AE000452	3326	3106	b3750, b3751	rbsC, rpsB

EXAMPLE 4

Identification of Genes and their Corresponding Operons Affected by Antisense Inhibition

The sequencing of the entire *E. coli* genome is described in Blattner et al., Science 277:1453-1474(1997) the entirety of which is hereby incorporated by reference and the sequence of the genome is listed in GenBank Accession No.U00096, the disclosure of which is incorporated herein by reference in its entirety. The operons to which the proliferation-inhibiting nucleic acids correspond were identified using RegulonDB and information in the literature. The coordinates of the boundaries of these operons on the *E. coli* genome are listed in Table III. Table II lists the molecule numbers of the inserts

containing the growth inhibiting nucleic acid fragments, the genes in the operons corresponding to the inserts, the SEQ ID NOs of the genes containing the inserts, the SEQ ID NOs of the proteins encoded by the genes, the start and stop points of the genes on the *E. coli* genome, the orientation of the genes on the genome, whether the operons are predicted or documented, and the predicted functions of the genes. The identified operons, their putative functions, and whether or not the genes are presently thought to be required for proliferation are discussed below.

Functions for the identified genes were determined by using either Blattner functional class designations or by comparing identified sequence with known sequences in various databases. A variety of biological functions were noted for the genes to which the clones of the present invention correspond. The functions for the genes of interest appear in Table II.

The proteins that are listed in Table II are involved in a wide range of biological functions.

TABLE II
All Operon Data with Whole Chromosome Coordinates

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA118a-d	*rpmJ*	94	287	3440255	3440371	<	documented	50S ribosomal subunit protein X (L36)
	prlA	95	288	3440403	3441734	<		*SecY*: multispanning membrane protein, translocator of proteins
	rplO	96	289	3441742	3442176	<		50S ribosomal subunit protein L15
	rpmD	97	290	3442180	3442359	<		50S ribosomal subunit protein L30
	rpsE	98	291	3442363	3442866	<		eps, spc, spcA; 30S ribosomal subunit protein S5
	rplR	99	292	3442881	3443234	<		50S ribosomal subunit protein L18
	rplF	100	293	3443244	3443777	<		50S ribosomal subunit protein L6; gentamicin sensitivity
	rpsH	101	294	3443790	3444182	<		30S ribosomal subunit protein S8
	rpsN	102	295	3444216	3444521	<		30S ribosomal subunit protein S14
	rplE	103	296	3444536	3445075	<		50S ribosomal subunit protein L5
	rplX	104	297	3445090	3445404	<		50S ribosomal subunit protein L24
	rplN	105	298	3445415	3445786	<		50S ribosomal subunit protein L14
EcXA119	*b2882*	106	299	3022315	3023772	>	predicted	Unknown
	b2883	107	300	3023787	3025106	>		Unknown
	b2884	108	301	3025142	3025711	>		Unknown
	b2885	109	302	3025678	3026508	>		Unknown
EcXA120	*b1509*	110	303	1590689	1592089	<	predicted	Unknown
	ydeK	111	304	1592133	1596110	<		Unknown
EcXA121	*yjfQ*	112	305	4415276	4416031	<	predicted	Unknown
EcXA122	*b2106*	113	306	2183937	2184761	>	predicted	Unknown
	b2107	114	307	2184800	2185318	>		Unknown
	yehA	115	308	2185400	2186434	<	predicted	Unknown
	yehB	116	309	2186450	2188930	<		Unknown
	yehC	117	310	2188946	2189665	<		Unknown
	yehD	118	311	2189700	2190242	<		Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA123	*yheL*	119	312	3472315	3472602	<	predicted	Unknown
	b3344	120	313	3472610	3472969	<		Unknown
	yheN	121	314	3472969	3473355	<		Unknown
	b3346	122	315	3473355	3474089	<		Unknown
EcXA124	*rpsQ*	123	316	3445951	3446205	<	documented	neaA; 30S ribosomal subunit protein S17
	rpmC	124	317	3446205	3446396	<		50S ribosomal subunit protein L29
	rplP	125	318	3446396	3446806	<		50S ribosomal subunit protein L16
	rpsC	126	319	3446819	3447520	<		30S ribosomal subunit protein S3
	rplV	127	320	3447538	3447870	<		eryB; erythromycin sensitivity; 50S ribosomal subunit protein L22
	rpsS	128	321	3447885	3448163	<		30S ribosomal subunit protein S19
	rplB	129	322	3448180	3449001	<		50S ribosomal subunit protein L2
	rplW	130	323	3449019	3449321	<		50S ribosomal subunit protein L23
	rplD	131	324	3449318	3449923	<		eryA; 50S ribosomal subunit protein L4; erythromycin sensitivity
	rplC	132	325	3449934	3450563	<		50S ribosomal subunit protein L3
	rpsJ	133	326	3450596	3450907	<		nusE; 30S ribosomal subunit protein S10
EcXA125	*yiiL*	134	327	4090705	4091019	<	predicted	Unknown
EcXA126	*yicC*	135	328	3814303	3815166	>	predicted	Unknown
EcXA127a-b	*b2851*	136	329	2989290	2989781	>	predicted	Unknown
EcXA128	*nagE*	137	330	703167	705113	>	documented	pstN; N-acetylglucosamine-specific enzyme II of phosphotransferase system
EcXA129	*yadC*	138	331	149715	150953	<	predicted	Unknown
	yadK	139	332	151003	151599	<		Unknown
	yadL	140	333	151626	152231	<		Unknown
	yadM	141	334	152243	152854	<		Unknown
EcXA130	*atpC*	142	335	3913181	3913600	<	documented	papG, uncC; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpD	143	336	3913621	3915003	<		papB, uncD; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpG	144	337	3915030	3915893	<		papC, uncG; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	atpA	145	338	3915944	3917485	<		papA, uncA; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpH	146	339	3917498	3918031	<		papE, uncH; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpF	147	340	3918046	3918516	<		papF, uncF; membrane-bound ATP synthase, F0 sector, subunit b (EC 3.6.1.3)
	atpE	148	341	3918578	3918817	<		papH, uncE; membrane-bound ATP synthase, F0 sector, subunit c; DCCD (EC 3.6.1.3)
	atpB	149	342	3918864	3919679	<		papD, uncB; membrane-bound ATP synthase, F0 sector, subunit a (EC 3.6.1.3)
	atpI	150	343	3919688	3920080	<		uncI; membrane-bound ATP synthase subunit, F1-F0-type proton-ATPase (EC 3.6.1.34)
EcXA131	*b3015*	151	344	3156944	3158185	<	predicted	Unknown
	b3016	152	345	3158185	3159162	<		Unknown
EcXA132	*narV*	153	346	1533961	1534641	<	documented	regulationCryptic nitrate reductase II, -subunit
	narW	154	347	1534638	1535333	<		regulationCryptic nitrate reductase II, -subunit
	narY	155	348	1535333	1536877	<		regulationCryptic NR II, -subunit
	narZ	156	349	1536874	1540614	<		regulationCryptic NR II, -subunit
EcXA133a-c							Same operon as EcXA118	
EcXA134	*sspB*	157	350	3373914	3374411	<	documented	Stress response protein
	sspA	158	351	3374417	3375055	<		pog; stress response protein
EcXA136a-b	*lysS*	159	352	3031677	3033194	<	documented	asuD herC; lysyl tRNA synthetase, constitutive
	prfB	160	353	3033204	3034302	<		supK; peptide chain release factor 2
	recJ	161	354	3034393	3036126	<	predicted	Single-stranded DNA-specific exonuclease, 5 -3

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	dsbC	162	355	3036132	3036842	<		xprA; periplasmic disulfide oxidoreductase, protein disulfide isomerase
	xerD	163	356	3036867	3037763	<		xprB; recombinase, site specific
EcXA137	*proB*	164	357	259612	260715	>	documented	pro2; -glutamyl kinase (EC 2.7.2.11)
	proA	165	358	260727	261980	>		pro1; -glutamyl phosphate reductase (EC 1.2.1.41)
EcXA138	*yhaB*	166	359	3265474	3266034	>	predicted	
	yhaC	167	360	3266056	3267243	>		
EcXA139a-b	*b0874*	168	361	913181	914128	<	predicted	
EcXA140a-b	*ilvN*	169	362	3848429	3848719	<	documented	Acetohydroxy acid synthase I, small subunit (EC 4.1.3.18); valine sensitive small subunit;positive regulator for thr and ilv
	ilvB	170	363	3848723	3850411	<		Acetolactate synthase I, valine sensitive (EC 4.1.3.18)
	b3672	171	364	3850517	3850615	<		
EcXA141	*yi52_7*	172	365	2099917	2100933	<	predicted	nucleotide sequence and phylogenetic relationships of a new E. coli insertion element. Schwartz E
	yefJ	173	366	2100938	2101411	<		Unknown
	yefI	174	367	2101413	2102531	<		Unknown
	yefH	175	368	2102516	2103106	<		Unknown
	yefG	176	369	2103087	2104079	<		Unknown
	rfc	177	370	2104082	2105248	<		O-antigen polymerase
	yefE	178	371	2105248	2106351	<		Unknown
EcXA142							Same operon as EcXA118	
EcXA143	*b2846*	179	372	2985498	2986190	>	predicted	Unknown
EcXA144a-c	*manA*	180	373	1686600	1687775	>	predicted	pmi; mannosephosphate isomerase (EC 5.3.1.8)
EcXA145	*thyA*	181	374	2962383	2963177	<	documented	Aminopterin, trimethoprim resistance; thymidylate synthetase (EC 2.1.1.45)

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	lgt	182	375	2963184	2964059	<		transferaseumpA; phosphatidylglycerol:prolipoprotein diacylglycerol transferase
EcXA146	*phoU*	183	376	3904481	3905206	<	documented	phoT; P uptake, high-affinity P-specific transport system, regulatory gene
	pstB	184	377	3905221	3905994	<		phoT; high-affinity P-specific transport; cytoplasmic ATP-binding protein
	pstA	185	378	3906177	3907067	<		R2pho, phoR2b, phoT; high-affinity P-specific transport
	pstC	186	379	3907067	3908026	<		phoW; high-affinity P-specific transport; cytoplasmic membrane component
	pstS	187	380	3908113	3909153	<		phoR2, nmpA, phoR2a, phoS, R2pho; high-affinity P-specific transport; periplasmic Pbinding
EcXA147	*b1111*	188	381	1167423	1168133	<	predicted	Unknown
EcXA148	*yifH*	189	382	3972208	3972753	>	predicted	Unknown
	yifI	190	383	3972758	3973888	>		Unknown
	yifJ	191	384	3973890	3975140	>		Unknown
	rffT	192	385	3976214	3977566	>		Synthesis of enterobacterial common antigen; Fuc4NAc transferase
	rffM	193	386	3977569	3978309	>		UDP-ManNAcA transferase
	b4404	194	387	3975137	3975361	>		Unknown
	b4405	195	388	3975603	3976217	>		Unknown
EcXA149a-b	*ygfA*	196	389	3054261	3054809	>	predicted	Unknown
EcXA150	*rpoZ*	197	390	3819733	3820008	>	documented	spoS; RNA polymerase,
	spoT	198	391	3820027	3822135	>		Guanosine 5 -diphosphate, 3 -diphosphate pyrophosphatase, ppGpp synthetase II activity
	spoU	199	392	3822142	3822831	>		
	recG	200	393	3822837	3824918	>		spoV? branch migration of Holliday junctions, junction-specific DNA helicase (see ruvABC)

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA151	*dctA*	201	394	3679791	3681077	<	predicted	Uptake of C-4 dicarboxylic acids; 3-fluoromalate resistance, D-tartrate resistant
EcXA152	*bcr*	202	395	2276590	2277780	<	predicted	bicA, bicR, sur, suxA; transmembrane; affects sulfathiazole-sulfonamide resistance
	yejD	203	396	2277808	2278503	<		Unknown
EcXA153	*yjbI*	204	397	4248534	4249862	>	predicted	Unknown
EcXA154	*b1983*	205	398	2054880	2055596	>	predicted	Unknown
EcXA155	*sseB*	206	399	2652177	2652962	<	predicted	Enhances serine sensitivity (inhibits homoserine deHase) on lactate; weaker than sseA
EcXA156	*yihN*	207	400	4059826	4061091	>	predicted	Unknown
EcXA157							Same operon as EcXA124	Unknown
EcXA158	*b2519*	208	401	2643033	2645345	<	predicted	Unknown
	b2520	209	402	2645346	2650307	<		Unknown
EcXA159	*b0482*	210	403	506510	507304	<	predicted	Unknown
EcXA160	*b1744*	211	404	1823979	1824947	<	predicted	Unknown
	b1745	212	405	1824940	1826283	<		Unknown
	b1746	213	406	1826280	1827758	<		Unknown
	b1747	214	407	1827755	1828789	<		Unknown
	b1748	215	408	1828786	1830006	<		Unknown
EcXA161	*sbcC*	216	409	411831	414977	<	predicted	Cosuppressor with sbcB of recB recC mutations
	sbcD	217	410	414974	416176	<		Cosuppressor with sbcB of recB recC mutations
EcXA162	*yjgU*	218	411	4490155	4490919	<	predicted	Unknown
	yjgV	219	412	4490943	4491974	<		Unknown
EcXA163	*yidC*	220	413	3882705	3884351	>	predicted	Unknown
EcXA164	*glpA*	221	414	2350667	2352295	>	documented	Glycerol-3-phosphate dehydrogenase (anaerobic) large subunit (EC 1.1.99.5)

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	glpB	222	415	2352285	2353544	>		sn-Glycerol-3-phosphate dehydrogenase (anaerobic) subunit (EC 1.1.99.5); membraneanchor
	glpC	223	416	2353541	2354731	>		sn-Glycerol-3-phosphate dehydrogenase (anaerobic) small subunit (EC 1.1.99.5)
EcXA165	*ybeJ*	224	417	686062	686970	<	predicted	Unknown
EcXA166	*yihE*	225	418	4039996	4040982	>	predicted	Unknown
	dsbA	226	419	4040999	4041625	>		iarA, ppfA; disulfide oxidoreductase, periplasmic protein disulfide-isomerase; role incytochrome c synthesis (EC 5.3.4.1)
EcXA167	*gatR_2*	227	420	2169417	2169755	<	documented	
	gatD	228	421	2169855	2170895	<		Galactitol-1-phosphate dehydrogenase
	gatC	229	422	2170943	2172298	<		Galactitol-specific enzyme IIC of PTS
	gatB	230	423	2172302	2172586	<		Galactitol-specific enzyme IIB of PTS
	gatA	231	424	2172617	2173069	<		Galactitol-specific enzyme IIA of phosphotransferase system (PTS)
	gatZ	232	425	2173079	2174341	<		Function unknown
	gatY	233	426	2174370	2175230	<		D-Tagatose-1,6-bisphosphate aldolase
EcXA168	*b1720*	234	427	1801118	1801591	>	predicted	Unknown
	b1721	235	428	1801602	1803017	>		Unknown
EcXA169							Same operon as EcXA118	Unknown
EcXA170	*b0362*	236	429	381728	382114	<	predicted	Unknown
	b0363	237	430	381963	383159	<		Unknown
EcXA171	*b0876*	238	431	915696	917354	>	predicted	Unknown
EcXA172	*b2345*	239	432	2461032	2462090	>	predicted	Unknown
EcXA173	*yhhJ*	240	433	3623310	3624437	<	predicted	Unknown
	b3486	241	434	3624434	3627118	<		Unknown
	yhiI	242	435	3627165	3628232	<		Unknown
EcXA174	*endA*	243	436	3088366	3089073	>	predicted	DNA-specific endonuclease I; extensive DNA breakdown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA175	*pykA*	244	437	1935673	1937115	>	predicted	Pyruvate kinase A (II); (EC 2.7.1.40)
EcXA176	*yjdG*	245	438	4346893	4347612	<	predicted	Unknown
	yjdH	246	439	4347609	4349240	<		Unknown
EcXA177	*b2378*	247	440	2493599	2494585	>	predicted	Unknown
EcXA178	*rpsT*	248	441	20815	21078	<	predicted	sups20; 30S ribosomal subunit protein S20
EcXA180	*yjcF*	249	442	4279362	4280654	<	predicted	Unknown
EcXA181	*ppc*	250	443	4148026	4150677	<	predicted	asp, glu; phosphoenolpyruvate carboxylase (EC 4.1.1.31)
EcXA182	*tktA*	251	444	3077663	3079654	<	predicted	Transketolase (EC 2.2.1.1)
EcXA183	*yabB*	252	445	89634	90092	>	predicted	Unknown
	yabC	253	446	90094	91035	>		Unknown
	ftsL	254	447	91032	91397	>		sensitiveCell division and growth; essential gene; cytoplasmic membrane protein
	ftsI	255	448	91413	93179	>		sensitivepbpB, sep; peptidoglycan synthetase; penicillin-binding protein 3
	murE	256	449	93166	94653	>		meso-Diaminopimelate adding enzyme
	murF	257	450	94650	96008	>		mra; D-alanyl:D-alanine adding enzyme
	mraY	258	451	96002	97084	>		UDP-N-acetylmuramoyl-pentapeptide:undecaprenyl-PO4 phosphatase (EC 2.7.8.13)
	murD	259	452	97087	98403	>		UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase (EC 6.3.2.9)
	ftsW	260	453	98403	99647	>		sensitiveCytoplasmic membrane required for PBP 2 expression; homology to rodA
	murG	261	454	99644	100711	>		UDP-NAc-glucosamine: NAc-muramyl-(pentapeptide) pyrophosphoryl-undecaprenolNAc-glucosamine transferase
	murC	262	455	100765	102240	>		L-Alanine adding enzyme
	ddlB	263	456	102233	103153	>		ddl; D-Alanine:D-alanine ligase
EcXA184	*cysC*	264	457	2871410	2872015	<	documented	Adenylylsulfate kinase (EC 2.7.1.25)
	cysN	265	458	2872015	2873442	<		ATP sulfurylase (ATP:sulfate adenylyltransferase)

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA185	*rplQ*	266	459	3437253	3437636	<	documented	50S ribosomal subunit protein L17
	rpoA	267	460	3437677	3438666	<		phs, sez; phage P2 vir1 resistance; RNA polymerase, -subunit (EC 2.7.7.6)
	rpsD	268	461	3438692	3439312	<		ramA, sud2; 30S ribosomal subunit protein S4
	rpsK	269	462	3439346	3439735	<		30S ribosomal subunit protein S11
	rpsM	270	463	3439752	3440108	<		30S ribosomal subunit protein S13
EcXA186	*yagG*	271	464	284619	286001	>	predicted	Unknown
	b0271	272	465	286013	287623	>		Unknown
EcXA187	*b2271*	273	466	2383874	2384851	>	predicted	Unknown
EcXA188	*b1113*	274	467	1168635	1169597	<	predicted	Unknown
EcXA189a-c	*b1828*	275	468	1908189	1909673	>	predicted	Unknown
EcXA190a-b	*b3052*	276	469	3192961	3194394	<	predicted	Unknown
	glnE	277	470	3194442	3197282	<		GS adenylyl transferase (EC 2.7.7.42)
	ygiF	278	471	3197305	3198606	<		Unknown
EcXA191a-b	*gidB*	279	472	3920685	3921308	<	predicted	Glucose effects on cell division, perhaps replication
	gidA	280	473	3921372	3923261	<		Glucose effects on cell division, perhaps replication
	mioC	281	474	3923640	3924083	<		Initiation of replication; transcription of 16-kDa protein proceeds through ori
EcXA192	*rbsD*	282	475	3930941	3931396	>	documented	rbsP; D-ribose high-affinity transport system
	rbsA	283	476	3931404	3932909	>		rbsP, rbsT; D-ribose high-affinity transport system (may have chemotaxis function)
	rbsC	284	477	3932914	3933879	>		rbsP, rbsT; D-ribose high-affinity transport system
	rbsB	285	478	3933904	3934794	>		prlB, rbsP; D-ribose periplasmic binding protein
	rbsK	286	479	3934920	3935849	>		Ribokinase (EC 2.7.1.15)

Functions for the identified genes were assigned using either Blattner functional class designations, functions referenced in Berlyn, MKB "Linkage Map of *Escherichia coli* K-12, Edition 10: The Traditional Map". (1998) Microbiol. Mol. Biol. Rev. September 62 (3): 814-984, or by comparing identified sequence with known sequences in various databases. A variety of biological functions were noted for the genes to which the clones of the present invention correspond. Biological functions for genes that lie on the same operon as an identified gene have also been made. The functions for the genes of interest appear in Table II.

The genes of interest have a variety of biological functions. For example, genes that are thought to function as transport or binding proteins, that participate in translation or post-translational modification, that are involved in carbon compound catabolism, that are thought to be enzymes, participate in cell processes, energy metabolism and biosynthetic functions appear in Table II. Genes that are involved in cell structure, transcription, RNA processing and degradation also appear in Table II.

Several of the expression vectors contain fragments that correspond to genes of unknown function or if the function is known, it is not known whether the gene is essential. For example, EcXA119, 120, 121, 122a-d, 123, 125, 126, 127a-b, 128, 129, 131, 132, 138, 139a-b, 141, 143, 146, 147, 14, 149a-b, 152, 153, 154, 155, 156, 158, 159, 160, 162, 163, 164, 165, 166, 167, 168, 170, 171, 172, 173, 176, 177, 180, 181, 186, 187, 188, 189a-b, 190a-b, 191a-b, and 192 are all exogenous nucleic acid sequences that correspond to *E. coli* proteins that have no known function or where the function has not been shown to be essential or nonessential.

The present invention reports a number of novel *E. coli* genes and operons that are required for proliferation. From the list of clone sequences identified here, each was identified to be a portion of a gene in an operon required for the proliferation of *E. coli*. Cloned sequences corresponding to genes already known to be required for proliferation in *E. coli* include EcXA118a-d, 124, 130, 133a-c, 136a-b, 142, 145, 150, 157, 169, 178, 182, 183 and 185 are exogenous nucleic acid sequences that correspond to *E. coli* genes that are known to be required for cellular proliferation. The remaining identified sequences correspond to *E. coli* genes previously undesignated as required for proliferation in the art.

An interesting observation of the present invention is that there are also several sequence fragments that correspond to *E. coli* genes that are not thought to be required for

E. coli proliferation. Nevertheless, under the conditions described above, the antisense expression of these gene fragments causes a reduction in cell growth. This result implies that the genes corresponding to the identified sequences are actually required for proliferation or are in operons required for proliferation. Molecule Nos. corresponding to these genes are EcXA128, 134, 137, 140a-b, 144a-c, 151, 161, 174, 175, and 184.

Following identification of the sequences of interest, these sequences were localized into operons. Since bacterial genes are expressed in a polycistronic manner, the antisense inhibition of a single gene in an operon might effect the expression of all the other genes on the operon or the genes down stream from the single gene identified. In order to determine which of the gene products in an operon are required for proliferation, each of the genes contained within an operon may be analyzed for their effect on viability as described below.

TABLE III

Operon Boundaries

Molecule Number	Start	Stop
EcXA118a-d	3440255	3445786
EcXA119	3022315	3026508
EcXA120	1590689	1596110
EcXA121	4415276	4416031
EcXA122	2183937	2190242
EcXA123	3472315	3474089
EcXA124	3445951	3450907
EcXA125	4090705	4091019
EcXA126	3814303	3815166
EcXA127a-b	2989290	2989781
EcXA128	703167	705113
EcXA129	149715	152854
EcXA130	3913181	3920080
EcXA131	3156944	3159162
EcXA132	1533961	1540614
EcXA133a-c	Same operon as EcXA118	
EcXA134	3373914	3375055
EcXA136a-b	3031677	3037763
EcXA137	259612	261980
EcXA138	3265474	3267243
EcXA139a-b	913181	914128
EcXA140a-b	3848429	3850615
EcXA141	2099917	2106351
EcXA142	Same operon as EcXA118	
EcXA143	2985498	2986190
EcXA144a-c	1686600	1687775
EcXA145	2962383	2964059
EcXA146	3904481	3909153
EcXA147	1167423	1168133
EcXA148	3972208	3976217
EcXA149a-b	3054261	3054809
EcXA150	3819733	3824918
EcXA151	3679791	3681077
EcXA152	2276590	2278503
EcXA153	4248534	4249862
EcXA154	2054880	2055596
EcXA155	2652177	2652962
EcXA156	4059826	4061091
EcXA157	Same operon as EcXA124	
EcXA158	2643033	2645345
	2645346	2650307

Molecule Number	Start	Stop
EcXA159	506510	507304
EcXA160	1823979	1830006
EcXA161	411831	416176
EcXA162	4490155	4491974
EcXA163	3882705	3884351
EcXA164	2350667	2354731
EcXA165	686062	686970
EcXA166	4039996	4041625
EcXA167	2169417	2175230
EcXA168	1801118	1803017
EcXA169	Same operon as EcXA118	
EcXA170	381728	383159
EcXA171	915696	917354
EcXA172	2461032	2462090
EcXA173	3623310	3628232
EcXA174	3088366	3089073
EcXA175	1935673	1937115
EcXA176	4346893	4349240
EcXA177	2493599	2494585
EcXA178	20815	21078
EcXA180	4279362	4280654
EcXA181	4148026	4150677
EcXA182	3077663	3079654
EcXA183	89634	103153
EcXA184	2871410	2873442
EcXA185	3437253	3440108
EcXA186	284619	287623
EcXA187	2383874	2384851
EcXA188	1168635	1169597
EcXA189a-c	1908189	1909673
EcXA190a-b	3192961	3198606
EcXA191a-b	3920685	3924083
EcXA192	3930941	3935849

EXAMPLE 5

Identification of Individual Genes within an Operon Required for Proliferation

The following example illustrates a method for determining which gene in an operon is required for proliferation. The clone insert corresponding to Molecule No. EcXA119 possesses nucleic acid sequence homology to the *E. coli* gene *b2883*. This gene is located in an operon containing the *b2882, b2883, b2884*, and *b2885* genes. To determine which gene or genes in this operon are required for proliferation, each gene is selectively inactivated using homologous recombination. Gene *b2885* is the first gene to be inactivated.

Deletion inactivation of a chromosomal copy of a gene in *E. coli* can be accomplished by integrative gene replacement. The principle of this method (Hamilton, C. M., et al 1989. *J. Bacteriol.* **171**: 4617-4622) is to construct a mutant allele of the targeted gene, introduce that allele into the chromosome using a conditional suicide vector, and then force the removal of the native wild type allele and vector sequences. This will replace the native gene with a desired mutation(s) but leave promoters, operators, etc. intact. Essentiality of a gene is determined either by deduction from genetic analysis or by conditional expression of a wild type copy of the targeted gene (trans complementation).

The first step is to generate a mutant *b2885* allele using PCR amplification. Two sets of PCR primers are chosen to produce a copy of *b2885* with a large central deletion to inactivate the gene. In order to eliminate polar effects, it is desirable to construct a mutant allele comprising an in-frame deletion of most or all of the coding region of the *b2885* gene. Each set of PCR primers is chosen such that a region flanking the gene to be amplified is sufficiently long to allow recombination (typically at least 500 nucleotides on each side of the deletion). The targeted deletion or mutation will be contained within this fragment. To facilitate cloning of the PCR product, the PCR primers may also contain restriction endonuclease sites found in the cloning region of a conditional knockout vector such as pKO3 (Link, et al 1997 *J. Bacteriol.* **179** (20): 6228-6237). Suitable sites include *Not*I, *Sal*I, *Bam*HI and *Sma*I. The b2885 gene fragments are produced using standard PCR conditions including, but not limited to, those outlined in the manufacturers directions for the Hot Start Taq PCR kit (Qiagen, Inc., Valencia, CA). The PCR reactions will produce two fragments that can be fused

together. Alternatively, crossover PCR can be used to generate a desired deletion in one step (Ho, S. N., et al 1989. *Gene* 77: 51-59, Horton, R. M., et al 1989. *Gene* 77: 61-68). The mutant allele thus produced is called a "null" allele because it cannot produce a functional gene product.

The mutant allele obtained from PCR amplification is cloned into the multiple cloning site of pKO3. Directional cloning of the *b2885* null allele is not necessary. The pKO3 vector has a temperature-sensitive origin of replication derived from pSC101. Therefore, clones are propagated at the permissive temperature of 30°C. The vector also contains two selectable marker genes: one that confers resistance to chloramphenicol and another, the *Bacillus subtilis sacB* gene, that allows for counter-selection on sucrose containing growth medium. Clones that contain vector DNA with the null allele inserted are confirmed by restriction endonuclease analysis and DNA sequence analysis of isolated plasmid DNA. The plasmid containing the *b2885* null allele insert is known as a knockout plasmid.

Once the knockout plasmid has been constructed and its sequence verified, it is transformed into a Rec^+ *E. coli* host cell. Transformation can be by any standard method such as electroporation. In some fraction of the transformed cells, plasmids will integrate into the *E. coli* chromosome by homologous recombination between the *b2885* null allele in the plasmid and the *b2885* gene in the chromosome. Transformant colonies in which such an event has occurred are readily selected by growth at the non-permissive temperature of 43°C and in the presence of choramphenicol. At this temperature, the plasmid will not replicate as an episome and will be lost from cells as they grow and divide. These cells are no longer resistant to chloramphenicol and will not grow when it is present. However, cells in which the knockout plasmid has integrated into the *E. coli* chromosome remain resistant to chloramphenicol and propagate.

Cells containing integrated knock-out plasmids are usually the result of a single crossover event that creates a tandem repeat of the mutant and native wild type alleles of *b2885* separated by the vector sequences. A consequence of this is that *b2885* will still be expressed in these cells. In order to determine if the gene is essential for growth, the wild type copy must be removed. This is accomplished by selecting for plasmid excision, a process in which homologous recombination between the two alleles results

in looping out of the plasmid sequences. Cells that have undergone such an excision event and have lost plasmid sequences including *sacB* gene are selected for by addition of sucrose to the medium. The *sacB* gene product converts sucrose to a toxic molecule. Thus counter selection with sucrose ensures that plasmid sequences are no longer present in the cell. Loss of plasmid sequences is further confirmed by testing for sensitivity to chloramphenicol (loss of the chloramphenicol resistance gene). The latter test is important because occasionally a mutation in the *sacB* gene can occur resulting in a loss of *sacB* function with no effect on plasmid replication (Link, et. al., 1997 *J. Bacteriol.* **179** (20): 6228-6237). These artifact clones retain plasmid sequences and are therefore still resistant to chloramphenicol.

In the process of plasmid excision, one of the two *b2885* alleles is lost from the chromosome along with the plasmid DNA. In general, it is equally likely that the null allele or the wild type allele will be lost. Therefore, if the *b2885* gene is not essential, half of the clones obtained in this experiment will have the wild type allele on the chromosome and half will have the null allele. However, if the *b2885* gene is essential, cells containing the null allele will not be obtained as a single copy of the null allele would be lethal.

To determine the essentiality of *b2885,* a statistically significant number of the resulting clones, at least 20, are analyzed by PCR amplification of the b2885 gene. Since the null allele is missing a significant portion of the *b2885* gene, its PCR product is significantly shorter than that of the wild type gene and the two are readily distinguished by gel electrophoretic analysis. The PCR products may also be subjected to sequence determination for further confirmation by methods well known to those in the art.

The above experiment is generally adequate for determining the essentiality of a gene such as *b2885.* However, it may be necessary or desirable to more directly confirm the essentiality of the gene. There are several methods by which this can be accomplished. In general, these involve three steps: 1) construction of an episome containing a wild type allele, 2) isolation of clones containing a single chromosomal copy of the mutant null allele as described above but in the presence of the episomal wild type allele, and then 3) determining if the cells survive when the expression of the episomal allele is shut off. In this case, the trans copy of wild type *b2885* is made by

PCR cloning of the entire coding region of *b2885* and inserting it in the sense orientation downstream of an inducible promoter such as the *E. coli lac* promoter. Transcription of this allele of *b2885* will be induced in the presence of IPTG which inactivates the *lac* repressor. Under IPTG induction *b2885* protein will be expressed as long as the recombinant gene also possesses a ribosomal binding site, also known as a "Shine-Dalgarno Sequence". The trans copy of *b2885* is cloned on a plasmid that is compatible with pSC101. Compatible vectors include p15A, pBR322, and the pUC plasmids, among others. Replication of the compatible plasmid will not be temperature-sensitive. The entire process of integrating the null allele of *b2885* and subsequent plasmid excision is carried out in the presence of IPTG to ensure the expression of functional *b2885* protein is maintained throughout. After the null *b2885* allele is confirmed as integrated on the chromosome in place of the wild type *b2885* allele, then IPTG is withdrawn and expression of functional *b2885* protein shut off. If the *b2885* gene is essential, cells will cease to proliferate under these conditions. However, if the *b2885* gene is not essential, cells will continue to proliferate under these conditions. In this experiment, essentiality is determined by conditional expression of a wild type copy of the gene rather than inability to obtain the intended chromosomal disruption.

An advantage of this method over some other gene disruption techniques is that the targeted gene can be deleted or mutated without the introduction of large segments of foreign DNA. Therefore, polar effects on downstream genes are eliminated or minimized. There are methods described to introduce inducible promoters upstream of potential essential bacterial genes. However in such cases, polarity from multiple transcription start points can be a problem. One way of preventing this is to insert a gene disruption cassette that contains strong transcriptional terminators upstream of the integrated inducible promoter (Zhang, Y, and Cronan, J. E. 1996 *J. Bacteriol.* **178** (12): 3614-3620). The described techniques will all be familiar to one of ordinary skill in the art.

Following the analysis of the *b2885* gene, the other genes of the operon are investigated to determine if they are required for proliferation.

EXAMPLE 6

Expression of the Proteins Encoded by Genes Identified as Required for *E. coli* Proliferation

The following is provided as one exemplary method to express the proliferation-required proteins encoded by the identified sequences described above. First, the initiation and termination codons for the gene are identified. If desired, methods for improving translation or expression of the protein are well known in the art. For example, if the nucleic acid encoding the polypeptide to be expressed lacks a methionine codon to serve as the initiation site, a strong Shine-Delgarno sequence, or a stop codon, these sequences can be added. Similarly, if the identified nucleic acid sequence lacks a transcription termination signal, this sequence can be added to the construct by, for example, splicing out such a sequence from an appropriate donor sequence. In addition, the coding sequence may be operably linked to a strong promoter or an inducible promoter if desired. The identified nucleic acid sequence or portion thereof encoding the polypeptide to be expressed is obtained by PCR from the bacterial expression vector or genome using oligonucleotide primers complementary to the identified nucleic acid sequence or portion thereof and containing restriction endonuclease sequences for *Nco*I incorporated into the 5' primer and *Bgl*II at the 5' end of the corresponding 3'-primer, taking care to ensure that the identified nucleic acid sequence is positioned in frame with the termination signal. The purified fragment obtained from the resulting PCR reaction is digested with *Nco*I and *Bgl*II, purified and ligated to an expression vector.

The ligated product is transformed into DH5α or some other *E. coli* strain suitable for the over expression of potential proteins. Transformation protocols are well known in the art. For example, transformation protocols are described in: **Current Protocols in Molecular Biology,** Vol. 1, Unit 1.8, (Ausubel, et al., Eds.) John Wiley & Sons, Inc. (1997). Positive transformants are selected after growing the transformed cells on plates containing 50-100 μg/ml Ampicillin (Sigma, St. Louis, Missouri). In one embodiment, the expressed protein is held in the cytoplasm of the host organism. In an alternate embodiment, the expressed protein is released into the culture medium. In still another alternative, the expressed protein can be sequestered in the periplasmic space and liberated therefrom using any one of a number of cell lysis techniques known in the art. For example, the osmotic shock cell lysis method described in Chapter 16 of **Current**

Protocols in Molecular Biology, Vol. 2, (Ausubel, et al., Eds.) John Wiley & Sons, Inc. (1997). Each of these procedures can be used to express a proliferation-required protein.

Expressed proteins, whether in the culture medium or liberated from the periplasmic space or the cytoplasm, are then purified or enriched from the supernatant using conventional techniques such as ammonium sulfate precipitation, standard chromatography, immunoprecipitation, immunochromatography, size exclusion chromatography, ion exchange chromatography, and HPLC. Alternatively, the secreted protein can be in a sufficiently enriched or pure state in the supernatant or growth media of the host to permit it to be used for its intended purpose without further enrichment. The purity of the protein product obtained can be assessed using techniques such as Coomassie or silver staining or using antibodies against the control protein. Coomassie and silver staining techniques are familiar to those skilled in the art.

Antibodies capable of specifically recognizing the protein of interest can be generated using synthetic peptides using methods well known in the art. See, **Antibodies: A Laboratory Manual,** (Harlow and Lane, Eds.) Cold Spring Harbor Laboratory (1988). For example, 15-mer peptides having a sequence encoded by the appropriate identified gene sequence of interest or portion thereof can be chemically synthesized. The synthetic peptides are injected into mice to generate antibodies to the polypeptide encoded by the identified nucleic acid sequence of interest or portion thereof. Alternatively, samples of the protein expressed from the expression vectors discussed above can be purified and subjected to amino acid sequencing analysis to confirm the identity of the recombinantly expressed protein and subsequently used to raise antibodies. An Example describing in detail the generation of monoclonal and polyclonal antibodies appears in Example 7.

The protein encoded by the identified nucleic acid sequence of interest or portion thereof can be purified using standard immunochromatography techniques. In such procedures, a solution containing the secreted protein, such as the culture medium or a cell extract, is applied to a column having antibodies against the secreted protein attached to the chromatography matrix. The secreted protein is allowed to bind the immunochromatography column. Thereafter, the column is washed to remove non-specifically bound proteins. The specifically bound secreted protein is then released from the column and recovered using standard techniques. These procedures are well known in the art.

In an alternative protein purification scheme, the identified nucleic acid sequence of interest or portion thereof can be incorporated into expression vectors designed for use in purification schemes employing chimeric polypeptides. In such strategies the coding sequence of the identified nucleic acid sequence of interest or portion thereof is inserted in-frame with the gene encoding the other half of the chimera. The other half of the chimera can be maltose binding protein (MBP) or a nickel binding polypeptide encoding sequence. A chromatography matrix having antibody to MBP or nickel attached thereto is then used to purify the chimeric protein. Protease cleavage sites can be engineered between the MBP gene or the nickel binding polypeptide and the identified expected gene of interest, or portion thereof. Thus, the two polypeptides of the chimera can be separated from one another by protease digestion.

One useful expression vector for generating maltose binding protein fusion proteins is pMAL (New England Biolabs), which encodes the *malE* gene. In the pMal protein fusion system, the cloned gene is inserted into a pMal vector downstream from the *malE* gene. This results in the expression of an MBP-fusion protein. The fusion protein is purified by affinity chromatography. These techniques as described are well known to those skilled in the art of molecular biology.

EXAMPLE 7

Production of an Antibody to an isolated *E. coli* Protein

Substantially pure protein or polypeptide is isolated from the transformed cells as described in Example 6. The concentration of protein in the final preparation is adjusted, for example, by concentration on a 10,000 molecular weight cut off AMICON filter device (Millipore, Bedford, MA), to the level of a few micrograms/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, G. and Milstein, C., **Nature 256**:495 (1975) or any of the well-known derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein or peptides derived therefrom over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as described by Engvall, E., "Enzyme immunoassay ELISA and EMIT," **Meth. Enzymol. 70**:419 (1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis, L. et al. **Basic Methods in Molecular Biology** Elsevier, New York. Section 21-2.

Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein or a peptide can be prepared by immunizing suitable animals with the expressed protein or peptides derived therefrom described above, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than larger molecules and can require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis, J. et al. **J. Clin. Endocrinol. Metab. 33**:988-991 (1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double

immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. et al., Chap. 19 in: **Handbook of Experimental Immunology** D. Wier (ed) Blackwell (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: **Manual of Clinical Immunology**, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D.C. (1980).

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. The antibodies can also be used in therapeutic compositions for killing bacterial cells expressing the protein.

EXAMPLE 8

Screening Chemical Libraries

A. Protein-Based Assays

Having isolated and expressed bacterial proteins shown to be required for bacterial proliferation, the present invention further contemplates the use of these expressed proteins in assays to screen libraries of compounds for potential drug candidates. The generation of chemical libraries is well known in the art. For example combinatorial chemistry can be used to generate a library of compounds to be screened in the assays described herein. A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining amino acids in every possible combination to yield peptides of a given length. Millions of chemical compounds theoretically can be synthesized through such combinatorial use of chemical building blocks. For example, one commentator observed that the systematic, combinatorial mixing of 100 interchangeable chemical building blocks results in the theoretical synthesis of 100 million tetrameric compounds or 10 billion pentameric compounds. (Gallop et al., "Applications of Combinatorial Technologies to Drug Discovery, Background and Peptide Combinatorial Libraries," **Journal of Medicinal Chemistry**, Vol. 37, No. 9, 1233-1250

(1994). Other chemical libraries known to those in the art may also be used, including natural product libraries.

Once generated, combinatorial libraries can be screened for compounds that possess desirable biological properties. For example, compounds which may be useful as drugs or to develop drugs would likely have the ability to bind to the target protein identified, expressed and purified as discussed above. Further, if the identified target protein is an enzyme, candidate compounds would likely interfere with the enzymatic properties of the target protein. Any enzyme can be a target protein. For example, the enzymatic function of a target protein can be to serve as a protease, nuclease, phosphatase, dehydrogenase, transporter protein, transcriptional enzyme, and any other type of enzyme known or unknown. Thus, the present invention contemplates using the protein products described above to screen combinatorial and other chemical libraries.

Those in the art will appreciate that a number of techniques exist for characterizing target proteins in order to identify molecules useful for the discovery and development of therapeutics. For example, some techniques involve the generation and use of small peptides to probe and analyze target proteins both biochemically and genetically in order to identify and develop drug leads. Such techniques include the methods described in PCT publications No. WO9935494, WO9819162, WO9954728, the disclosures of which are incorporated herein by reference in their entireties.

In another example, the target protein is a serine protease and the substrate of the enzyme is known. The present example is directed towards the analysis of libraries of compounds to identify compounds that function as inhibitors of the target enzyme. First, a library of small molecules is generated using methods of combinatorial library formation well known in the art. U.S. Patent NOs. 5,463,564 and 5,574, 656, to Agrafiotis, et al., entitled "System and Method of Automatically Generating Chemical Compound with Desired Properties," the disclsoures of which are incorporated herein by reference in their entireties, are two such teachings. Then the library compounds are screened to identify library compounds that possess desired structural and functional properties. U.S. Patent No. 5,684,711, the disclosure of which is incorporated herein by reference in its entirety, also discusses a method for screening libraries.

To illustrate the screening process, the combined target and chemical compounds of the library are exposed to and permitted to interact with the purified enzyme. A labeled

substrate is added to the incubation. The label on the substrate is such that a detectable signal is emitted from metabolized substrate molecules. The emission of this signal permits one to measure the effect of the combinatorial library compounds on the enzymatic activity of target enzymes. The characteristics of each library compound is encoded so that compounds demonstrating activity against the enzyme can be analyzed and features common to the various compounds identified can be isolated and combined into future iterations of libraries.

Once a library of compounds is screened, subsequent libraries are generated using those chemical building blocks that possess the features shown in the first round of screen to have activity against the target enzyme. Using this method, subsequent iterations of candidate compounds will possess more and more of those structural and functional features required to inhibit the function of the target enzyme, until a group of enzyme inhibitors with high specificity for the enzyme can be found. These compounds can then be further tested for their safety and efficacy as antibiotics for use in mammals.

It will be readily appreciated that this particular screening methodology is exemplary only. Other methods are well known to those skilled in the art. For example, a wide variety of screening techniques are known for a large number of naturally-occurring targets when the biochemical function of the target protein is known.

B. Cell-based Assays

Current cell-based assays used to identify or to characterize compounds for drug discovery and development frequently depend on detecting the ability of a test compound to inhibit the activity of a target molecule located within a cell or located on the surface of a cell. An advantage of cell-based assays is that they allow the effect of a compound on a target molecule's activity to be detected within the physiologically relevant environment of the cell as opposed to an *in vitro* environment. Most often such target molecules are proteins such as enzymes, receptors and the like. However, target molecules may also include other molecules such as DNAs, lipids, carbohydrates and RNAs including messenger RNAs, ribosomal RNAs, tRNAs and the like. A number of highly sensitive cell-based assay methods are available to those of skill in the art to detect binding and interaction of test compounds with specific target molecules. However, these methods are generally not highly effective when the test compound

binds to or otherwise interacts with its target molecule with moderate or low affinity. In addition, the target molecule may not be readily accessible to a test compound in solution, such as when the target molecule is located inside the cell or within a cellular compartment such as the periplasm of a bacterial cell. Thus, current cell-based assay methods are limited in that they are not effective in identifying or characterizing compounds that interact with their targets with moderate to low affinity or compounds that interact with targets that are not readily accessible.

Cell-based assay methods of the present invention have substantial advantages over current cell-based assays practiced in the art. These advantages derive from the use of sensitized cells in which the level or activity of a proliferation-required gene product (the target molecule) has been specifically reduced to the point where the presence or absence of its function becomes a rate-determining step for cellular proliferation. Bacterial, fungal, plant, or animal cells can all be used with the present method. Such sensitized cells become much more sensitive to compounds that are active against the affected target molecule. Thus, cell-based assays of the present invention are capable of detecting compounds exhibiting low or moderate potency against the target molecule of interest because such compounds are substantially more potent on sensitized cells than on non-sensitized cells. The affect may be such that a test compound may be two to several times more potent, at least 10 times more potent, at least 20 times more potent, at least 50 times more potent, at least 100 times more potent, at least 1000 times more potent, or even more than 1000 times more potent when tested on the sensitized cells as compared to the non-sensitized cells.

Due in part to the increased appearance of antibiotic resistance in pathogenic microorganisms and to the significant side-effects associated with some currently used antibiotics, novel antibiotics acting at new targets are highly sought after in the art. Yet, another limitation in the current art related to cell-based assays is the problem of identifying hits against the same kinds of target molecules in the same limited set of biological pathways over and over again. This may occur when compounds acting at such new targets are discarded, ignored or fail to be detected because compounds acting at the "old" targets are encountered more frequently and are more potent than compounds acting at the new targets. As a result, the majority of antibiotics in use

currently interact with a relatively small number of target molecules within an even more limited set of biological pathways.

The use of sensitized cells of the current invention provides a solution to the above problem in two ways. First, desired compounds acting at a target of interest, whether a new target or a previously known but poorly exploited target, can now be detected above the "noise" of compounds acting at the "old" targets due to the specific and substantial increase in potency of such desired compounds when tested on the sensitized cells of the current invention. Second, the methods used to sensitize cells to compounds acting at a target of interest may also sensitize these cells to compounds acting at other target molecules within the same biological pathway. For example, expression of an antisense molecule to a gene encoding a ribosomal protein is expected to sensitize the cell to compounds acting at that ribosomal protein and may also sensitize the cells to compounds acting at any of the ribosomal components (proteins or rRNA) or even to compounds acting at any target which is part of the protein synthesis pathway. Thus an important advantage of the present invention is the ability to reveal new targets and pathways that were previously not readily accessible to drug discovery methods.

Sensitized cells of the present invention are prepared by reducing the activity or level of a target molecule. The target molecule may be a gene product, such as an RNA or polypeptide produced from the proliferation-required nucleic acids described herein. Alternatively, the target may be a gene product such as an RNA or polypeptide which is produced from a sequence within the same operon as the proliferation-required nucleic acids described herein. In addition, the target may be an RNA or polypeptide in the same biological pathway as the proliferation-required nucleic acids described herein. Such biological pathways include, but are not limited to, enzymatic, biochemical and metabolic pathways as well as pathways involved in the production of cellular structures such the cell wall.

Current methods employed in the arts of medicinal and combinatorial chemistries are able to make use of structure-activity relationship information derived from testing compounds in various biological assays including direct binding assays and cell-based assays. Occasionally compounds are directly identified in such assays that are sufficiently potent to be developed as drugs. More often, initial hit compounds

exhibit moderate or low potency. Once a hit compound is identified with low or moderate potency, directed libraries of compounds are synthesized and tested in order to identify more potent leads. Generally these directed libraries are combinatorial chemical libraries consisting of compounds with structures related to the hit compound but containing systematic variations including additions, subtractions and substitutions of various structural features. When tested for activity against the target molecule, structural features are identified that either alone or in combination with other features enhance or reduce activity. This information is used to design subsequent directed libraries containing compounds with enhanced activity against the target molecule. After one or several iterations of this process, compounds with substantially increased activity against the target molecule are identified and may be further developed as drugs. This process is facilitated by use of the sensitized cells of the present invention since compounds acting at the selected targets exhibit increased potency in such cell-based assays, thus; more compounds can now be characterized providing more useful information than would be obtained otherwise.

Thus, it is now possible using cell-based assays of the present invention to identify or characterize compounds that previously would not have been readily identified or characterized including compounds that act at targets that previously were not readily exploited using cell-based assays. The process of evolving potent drug leads from initial hit compounds is also substantially improved by the cell-based assays of the present invention because, for the same number of test compounds, more structure-function relationship information is likely to be revealed.

The method of sensitizing a cell entails selecting a suitable gene or operon. A suitable gene or operon is one whose expression is required for the proliferation of the cell to be sensitized. The next step is to introduce into the cells to be sensitized, an antisense RNA capable of hybridizing to the suitable gene or operon or to the RNA encoded by the suitable gene or operon. Introduction of the antisense RNA can be in the form of an expression vector in which antisense RNA is produced under the control of an inducible promoter. The amount of antisense RNA produced is limited by varying the inducer concentration to which the cell is exposed and thereby varying the activity of the promoter driving transcription of the antisense RNA. Thus, cells are sensitized

by exposing them to an inducer concentration that results in a sub-lethal level of antisense RNA expression.

In one embodiment of the cell-based assays, the identified exogenous *E. coli* nucleotide sequences of the present invention are used to inhibit the production of a proliferation-required protein. Expression vectors producing antisense RNA complementary to identified genes required for proliferation are used to limit the concentration of a proliferation-required protein without severely inhibiting growth. To achieve that goal, a growth inhibition dose curve of inducer is calculated by plotting various doses of inducer against the corresponding growth inhibition caused by the antisense expression. From this curve, various percentages of antisense induced growth inhibition, from 1 to 100% can be determined. If the promoter contained in the expression vector contains a *lac* operator the transcription is regulated by *lac* repressor and expression from the promoer is inducible with IPTG. For example, the highest concentration of the inducer IPTG that does not reduce the growth rate significantly (0% growth inhibition) can be predicted from the curve. Cellular proliferation can be monitored by growth medium turbidity via OD measurements. In another example, the concentration of inducer that reduces growth by 25% can be predicted from the curve. In still another example, a concentration of inducer that reduces growth by 50% can be calculated. Additional parameters such as colony forming units (cfu) can be used to measure cellular viability.

Cells to be assayed are exposed to the above-determined concentrations of inducer. The presence of the inducer at this sub-lethal concentration reduces the amount of the proliferation required gene product to a low amount in the cell that will limit but not prevent growth. Cells grown in the presence of this concentration of inducer are therefore specifically more sensitive to inhibitors of the proliferation-required protein or RNA of interest or to inhibitors of proteins or RNAs in the same biological pathway as the proliferation-required protein or RNA of interest but not to inhibitors of unrelated proteins or RNAs.

Cells pretreated with sub-inhibitory concentrations of inducer and thus containing a reduced amount of proliferation-required target gene product are then used to screen for compounds that reduce cell growth. The sub-lethal concentration of inducer may be any concentration consistent with the intended use of the assay to identify candidate

compounds to which the cells are more sensitive. For example, the sub-lethal concentration of the inducer may be such that growth inhibition is at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60% at least about 75%, 90%, 95% or more. Cells which are pre-sensitized using the preceding method are more sensitive to inhibitors of the target protein because these cells contain less target protein to be inhibited than do wild-type cells.

In another embodiment of the cell-based assays of the present invention, the level or activity of a proliferation required gene product is reduced using a mutation, such as a temperature sensitive mutation, in the proliferation-required sequence and an antisense nucleic acid complementary to the proliferation-required sequence. Growing the cells at an intermediate temperature between the permissive and restrictive temperatures of the temperature sensitive mutant where the mutation is in a proliferation-required gene produces cells with reduced activity of the proliferation-required gene product. The antisense RNA complementary to the proliferation-required sequence further reduces the activity of the proliferation required gene product. Drugs that may not have been found using either the temperature sensitive mutation or the antisense nucleic acid alone may be identified by determining whether cells in which expression of the antisense nucleic acid has been induced and which are grown at a temperature between the permissive temperature and the restrictive temperature are substantially more sensitive to a test compound than cells in which expression of the antisense nucleic acid has not been induced and which are grown at a permissive temperature. Also drugs found previously from either the antisense nucleic acid alone or the temperature sensitive mutation alone may have a different sensitivity profile when used in cells combining the two approaches, and that sensitivity profile may indicate a more specific action of the drug in inhibiting one or more activities of the gene product.

Temperature sensitive mutations may be located at different sites within the gene and correspond to different domains of the protein. For example, the *dnaB* gene of *Escherichia coli* encodes the replication fork DNA helicase. DnaB has several domains, including domains for oligomerization, ATP hydrolysis, DNA binding, interaction with primase, interaction with DnaC, and interaction with DnaA [(Biswas,

E.E. and Biswas, S.B. 1999 Mechanism and DnaB helicase of *Escherichia coli*: structural domains involved in ATP hydrolysis, DNA binding, and oligomerization. Biochem. **38**:10919-10928; Hiasa, H. and Marians, K.J. 1999 Initiation of bidirectional replication at the chromosomal origin is directed by the interaction between helicase and primase. J. Biol. Chem. **274**:27244-27248; San Martin, C., Radermacher, M., Wolpensinger, B., Engel, A., Miles, C.S., Dixon, N.E., and Carazo, J.M. 1998 Three-dimensional reconstructions from cryoelectron microscopy images reveal an intimate complex between helicase DnaB and its loading partner DnaC. Structure **6**:501-9; Sutton, M.D., Carr, K.M., Vicente, M., and Kaguni, J.M. 1998 *Escherichia coli* DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal. J. Biol. Chem. **273**:34255-62), the disclosures of which are incorporated herein by reference in their entireties]. Temperature sensitive mutations in different domains of DnaB confer different phenotypes at the restrictive temperature, which include either an abrupt stop or slow stop in DNA replication with or without DNA breakdown (Wechsler, J.A. and Gross, J.D. 1971. *Escherichia coli* mutants temperature-sensitive for DNA synthesis. Mol. Gen. Genetics **113**:273-284, the disclosure of which is incorporated herein by reference in its entirety) and termination of growth or cell death. Combining the use of temperature sensitive mutations in the *dnaB* gene that cause cell death at the restrictive temperature with an antisense to the *dnaB* gene could lead to the discovery of very specific and effective inhibitors of one or a subset of activities exhibited by DnaB.

It will be appreciated that the above method may be performed with any mutation which reduces but does not eliminate the activity or level of the gene product which is required for proliferation.

When screening for antimicrobial agents against a gene product required for proliferation, growth inhibition of cells containing a limiting amount of that proliferation-required gene product can be assayed. Growth inhibition can be measured by directly comparing the amount of growth, measured by the optical density of the growth medium, between an experimental sample and a control sample. Alternative methods for assaying cell proliferation include measuring green fluorescent protein (GFP) reporter construct emissions, various enzymatic activity assays, and other methods well known in the art.

It will be appreciated that the above method may be performed in solid phase, liquid phase or a combination of the two. For example, cells grown on nutrient agar containing the inducer of the antisense construct may be exposed to compounds spotted onto the agar surface. A compound's effect may be judged from the diameter of the resulting killing zone, the area around the compound application point in which cells do not grow. Multiple compounds may be transferred to agar plates and simultaneously tested using automated and semi-automated equipment including but not restricted to multi-channel pipettes (for example the Beckman Multimek) and multi-channel spotters (for example the Genomic Solutions Flexys). In this way multiple plates and thousands to millions of compounds may be tested per day.

The compounds may also be tested entirely in liquid phase using microtiter plates as described below. Liquid phase screening may be performed in microtiter plates containing 96, 384, 1536 or more wells per microtiter plate to screen multiple plates and thousands to millions of compounds per day. Automated and semi-automated equipment may be used for addition of reagents (for example cells and compounds) and determination of cell density.

EXAMPLE 9

Cell-based Assay Using Antisense Complementary to Genes Encoding Ribosomal Proteins

The effectiveness of the above cell-based assay was validated using constructs expressing antisense RNA to the proliferation required *E. coli* genes *rplL*, *rplJ*, and *rplW* encoding ribosomal proteins L7/L12, L10 and L23 respectively. These proteins are part of the protein synthesis apparatus of the cell and as such are required for proliferation. These constructs were used to test the effect of antisense expression on cell sensitivity to antibiotics known to bind to the ribosome and thereby inhibit protein synthesis. Constructs expressing antisense RNA to several other genes (*elaD*, *visC*, *yohH*, and *atpE/B*), the products of which are not involved in protein synthesis were used for comparison.

First, pLEX5BA (Krause et al., J. Mol. Biol. 274: 365 (1997), the disclosure of which is incorporated herein by reference in its entirety) expression vectors containing antisense constructs to either *rplW* or to *elaD* were introduced into separate *E. coli* cell populations. Vector introduction is a technique well known to those of ordinary skill in

the art. The expression vectors of this example contain IPTG inducible promoters that drive the expression of the antisense RNA in the presence of the inducer. However, those skilled in the art will appreciate that other inducible promoters may also be used. Suitable expression vectors are also well known in the art. The *E. coli* antisense clones to genes encoding ribosomal proteins L7/L12, L10 and L23 were used to test the effect of antisense expression on cell sensitivity to the antibiotics known to bind to these proteins. Expression vectors containing antisense to either the genes encoding L7/L12 and L10 or L23 were introduced into separate *E. coli* cell populations.

The cell populations were exposed to a range of IPTG concentrations in liquid medium to obtain the growth inhibitory dose curve for each clone (Fig. 1). First, seed cultures were grown to a particular turbidity that is measured by the optical density (OD) of the growth solution. The OD of the solution is directly related to the number of bacterial cells contained therein. Subsequently, sixteen 200 ul liquid medium cultures were grown in a 96 well microtiter plate at 37° C with a range of IPTG concentrations in duplicate two-fold serial dilutions from 1600 uM to 12.5 uM (final concentration). Additionally, control cells were grown in duplicate without IPTG. These cultures were started from equal amounts of cells derived from the same initial seed culture of a clone of interest. The cells were grown for up to 15 hours and the extent of growth was determined by measuring the optical density of the cultures at 600 nm. When the control culture reached mid-log phase the percent growth (relative to the control culture) for each of the IPTG containing cultures was plotted against the log concentrations of IPTG to produce a growth inhibitory dose response curve for the IPTG. The concentration of IPTG that inhibits cell growth to 50% (IC_{50}) as compared to the 0 mM IPTG control (0% growth inhibition) was then calculated from the curve. Under these conditions, an amount of antisense RNA was produced that reduced the expression levels of *rplW* and *elaD* to a degree such that growth was inhibited by 50%.

Alternative methods of measuring growth are also contemplated. Examples of these methods include measurements of proteins, the expression of which is engineered into the cells being tested and can readily be measured. Examples of such proteins include green fluorescent protein (GFP) and various enzymes.

Cells were pretreated with the selected concentration of IPTG and then used to test the sensitivity of cell populations to tetracycline, erythromycin and other protein

synthesis inhibitors. Figure 2 is an IPTG dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing either an antisense clone to the *E. coli rplW* gene (AS-*rplW*) which encodes ribosomal protein L23 which is required for protein synthesis and essential for cell proliferation, or an antisense clone to the *elaD* (AS-*elaD*) gene which is not known to be involved in protein synthesis and which is also essential for proliferation.

An example of a tetracycline dose response curve is shown in Figures 2A and 2B for the *rplW* and *elaD* genes, respectively. Cells were grown to log phase and then diluted into media alone or media containing IPTG at concentrations which give 20% and 50% growth inhibition as determined by IPTG dose response curves. After 2.5 hours, the cells were diluted to a final OD_{600} of 0.002 into 96 well plates containing (1) +/- IPTG at the same concentrations used for the 2.5 hour pre-incubation; and (2) serial two-fold dilutions of tetracycline such that the final concentrations of tetracycline range from 1 μg/ml to 15.6 ng/ml and 0 μg/ml. The 96 well plates were incubated at 37°C and the OD_{600} was read by a plate reader every 5 minutes for up to 15 hours. For each IPTG concentration and the no IPTG control, tetracycline dose response curves were determined when the control (absence of tetracycline) reached 0.1 OD_{600}. To compare tetracycline sensitivity with and without IPTG, tetracycline IC_{50s} were determined from the dose response curves (Figs. 3A-B). Cells with reduced levels of L23 (AS-*rplW*) showed increased sensitivity to tetracycline (Fig. 2A) as compared to cells with reduced levels of the *elaD* gene product (AS-*elaD)* (Fig. 2B). Figure 3 shows a summary bar chart in which the ratios of tetracycline IC_{50s} determined in the presence of IPTG which gives 50% growth inhibition versus tetracycline IC_{50s} determined without IPTG (fold increase in tetracycline sensitivity) were plotted. Cells with reduced levels of either L7/L12 (encoded by genes *rplL*, *rplJ*) or L23 (encoded by the *rplW* gene) showed increased sensitivity to tetracycline (Fig. 3). Cells expressing antisense to genes not known to be involved in protein synthesis (AS-*atpB/E*, AS-*visC*, AS-*elaD*, AS-*yohH*) did not show the same increased sensitivity to tetracycline, validating the specificity of this assay (Fig. 3).

In addition to the above, it has been observed in initial experiments that clones expressing antisense RNA to genes involved in protein synthesis (including genes encoding ribosomal proteins L7/L12 & L10, L7/L12 alone, L22, and L18, as well as

genes encoding rRNA and Elongation Factor G) have increased sensitivity to the macrolide, erythromycin, whereas clones expressing antisense to the non-protein synthesis genes *elaD, atpB/E* and *visC* do not. Furthermore, the clone expressing antisense to *rplL* and *rplJ* does not show increased sensitivity to nalidixic acid and ofloxacin, antibiotics which do not inhibit protein synthesis.

The results with the ribosomal protein genes *rplL, rplJ*, and *rplW* as well as the initial results using various other antisense clones and antibiotics show that limiting the concentration of an antibiotic target makes cells more sensitive to the antimicrobial agents that specifically interact with that protein. The results also show that these cells are sensitized to antimicrobial agents that inhibit the overall function in which the protein target is involved but are not sensitized to antimicrobial agents that inhibit other functions.

The cell-based assay described above may also be used to identify the biological pathway in which a proliferation-required nucleic acid or its gene product lies. In such methods, cells expressing a sub-lethal level of antisense to a target proliferation-required nucleic acid and control cells in which expression of the antisense has not been induced are contacted with a panel of antibiotics known to act in various pathways. If the antibiotic acts in the pathway in which the target proliferation-required nucleic acid or its gene product lies, cells in which expression of the antisense has been induced will be more sensitive to the antibiotic than cells in which expression of the antisense has not been induced.

As a control, the results of the assay may be confirmed by contacting a panel of cells expressing antisense nucleic acids to many different proliferation-required genes including the target proliferation-required gene. If the antibiotic is acting specifically, heightened sensitivity to the antibiotic will be observed only in the cells expressing antisense to a target proliferation-required gene (or cells expressing antisense to other proliferation-required genes in the same pathway as the target proliferation-required gene) but will not be observed generally in all cells expressing antisense to proliferation-required genes.

Similarly, the above method may be used to determine the pathway on which a test compound, such as a test antibiotic acts. A panel of cells, each of which expresses antisense to a proliferation-required nucleic acid in a known pathway, is contacted with

a compound for which it is desired to determine the pathway on which it acts. The sensitivity of the panel of cells to the test compound is determined in cells in which expression of the antisense has been induced and in control cells in which expression of the antisense has not been induced. If the test compound acts on the pathway on which an antisense nucleic acid acts, cells in which expression of the antisense has been induced will be more sensitive to the compound than cells in which expression of the antisense has not been induced. In addition, control cells in which expression of antisense to proliferation-required genes in other pathways has been induced will not exhibit heightened sensitivity to the compound. In this way, the pathway on which the test compound acts may be determined.

The Example below provides one method for performing such assays.

EXAMPLE 10

Identification of the Pathway in which a Proliferation-Required Gene Lies or the Pathway on which an Antibiotic Acts

A. Preparation of Bacterial Stocks for Assay

To provide a consistent source of cells to screen, frozen stocks of host bacteria containing the desired antisense construct are prepared using standard microbiological techniques. For example, a single clone of the organism can be isolated by streaking out a sample of the original stock onto an agar plate containing nutrients for cell growth and an antibiotic for which the antisense construct contains a gene which confers resistance. After overnight growth an isolated colony is picked from the plate with a sterile needle and transferred to an appropriate liquid growth media containing the antibiotic required for maintenance of the plasmid. The cells are incubated at 30°C to 37°C with vigorous shaking for 4 to 6 hours to yield a culture in exponential growth. Sterile glycerol is added to 15% (volume to volume) and 100μL to 500 μL aliquots are distributed into sterile cryotubes, snap frozen in liquid nitrogen, and stored at -80°C for future assays.

B. Growth of Bacteria for Use in the Assay

A day prior to an assay, a stock vial is removed from the freezer, rapidly thawed (37°C water bath) and a loop of culture is streaked out on an agar plate containing nutrients for cell growth and an antibiotic to which the antisense construct confers resistance. After overnight growth at 37°C, ten randomly chosen, isolated colonies are transferred from the plate (sterile inoculum loop) to a sterile tube containing 5 mL of

LB medium containing the antibiotic to which the antisense vector confers resistance. After vigorous mixing to form a homogeneous cell suspension, the optical density of the suspension is measured at 600 nm (OD_{600}) and if necessary an aliquot of the suspension is diluted into a second tube of 5 mL, sterile, LB medium plus antibiotic to achieve an $OD_{600} \leq 0.02$ absorbance units. The culture is then incubated at 37° C for 1-2 hrs with shaking until the OD_{600} reaches OD 0.2 – 0.3. At this point the cells are ready to be used in the assay.

C. Selection of Media to be Used in Assay

Two fold dilution series of the inducer are generated in culture media containing the appropriate antibiotic for maintenance of the antisense construct. Several media are tested side by side and three to four wells are used to evaluate the effects of the inducer at each concentration in each media. For example, M9 minimal media, LB broth, TBD broth and Muller-Hinton media may be tested with the inducer IPTG at the following concentrations, 50 μM, 100 μM, 200 μM, 400 μM, 600 μM, 800 μM and 1000 μM. Equal volumes of test media-inducer and cells are added to the wells of a 384 well microtiter plate and mixed. The cells are prepared as described above and diluted 1:100 in the appropriate media containing the test antibiotic immediately prior to addition to the microtiter plate wells. For a control, cells are also added to several wells of each media that do not contain inducer, for example 0 μM IPTG. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of inducer is calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without inducer. The medium yielding greatest sensitivity to inducer is selected for use in the assays described below.

D. Measurement of Test Antibiotic Sensitivity in the Absence of Antisense Construct Induction

Two-fold dilution series of antibiotics with a known mechanism of action are generated in the culture media selected for further assay development that has been supplemented with the antibiotic used to maintain the construct. A panel of test antibiotics known to act on different pathways is tested side by side with three to four wells being used to evaluate the effect of a test antibiotic on cell growth at each concentration. Equal volumes of test antibiotic and cells are added to the wells of a 384

well microtiter plate and mixed. Cells are prepared as described above using the media selected for assay development supplemented with the antibiotic required to maintain the antisense construct and are diluted 1:100 in identical media immediately prior to addition to the microtiter plate wells. For a control, cells are also added to several wells that contain the solvent used to dissolve the antibiotics but no antibiotic. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of antibiotic is calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without antibiotic. A plot of percent inhibition against log[antibiotic concentration] allows extrapolation of an IC_{50} value for each antibiotic.

E. Measurement of Test Antibiotic Sensitivity in the Presence of Antisense Construct Inducer

The culture media selected for use in the assay is supplemented with inducer at concentrations shown to inhibit cell growth by 50% and 80% as described above and the antibiotic used to maintain the construct. Two fold dilution series of the panel of test antibiotics used above are generated in each of these media. Several antibiotics are tested side by side with three to four wells being used to evaluate the effects of an antibiotic on cell growth at each concentration, in each media. Equal volumes of test antibiotic and cells are added to the wells of a 384 well microtiter plate and mixed. Cells are prepared as described above using the media selected for use in the assay supplemented with the antibiotic required to maintain the antisense construct. The cells are diluted 1:100 into two 50 mL aliquots of identical media containing concentrations of inducer that have been shown to inhibit cell growth by 50% and 80 % respectively and incubated at 37°C with shaking for 2.5 hours. Immediately prior to addition to the microtiter plate wells, the cultures are adjusted to an appropriate OD_{600} (typically 0.002) by dilution into warm (37°C) sterile media supplemented with identical concentrations of the inducer and antibiotic used to maintain the antisense construct. For a control, cells are also added to several wells that contain solvent used to dissolve test antibiotics but which contain no antibiotic. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of antibiotic is

calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without antibiotic. A plot of percent inhibition against log[antibiotic concentration] allows extrapolation of an IC_{50} value for each antibiotic.

F. Determining the Specificity of the Test Antibiotics

A comparison of the IC_{50}s generated by antibiotics of known mechanism of action under antisense induced and non-induced conditions allows the pathway in which a proliferation-required nucleic acid lies to be identified. If cells expressing an antisense nucleic acid against a proliferation-required gene are selectively sensitive to an antibiotic acting via a particular pathway, then the gene against which the antisense acts is involved in the pathway in which the antibiotic acts.

G. Identification of Pathway in which a Test Antibiotic Acts

As discussed above, the cell-based assay may also be used to determine the pathway against which a test antibiotic acts. In such an analysis, the pathways against which each member of a panel of antisense nucleic acids acts are identified as described above. A panel of cells, each containing an inducible nucleic acid complementary to a gene in a known proliferation-required pathway, is contacted with a test antibiotic for which it is desired to determine the pathway on which it acts under inducing an non-inducing conditions. If heightened sensitivity is observed in induced cells expressing antisense complementary to a gene in a particular pathway but not in induced cells expressing antisense complementary to genes in other pathways, then the test antibiotic acts against the pathway for which heightened sensitivity was observed.

One skilled in the art will appreciate that further optimization of the assay conditions, such as the concentration of inducer used to induce antisense expression and/or the growth conditions used for the assay (for example incubation temperature and media components) may further increase the selectivity and/or magnitude of the antibiotic sensitization exhibited.

The following example confirms the effectiveness of the methods described above.

EXAMPLE 11

Identification of the Pathway in which a Proliferation-Required Gene Lies

Antibiotics of various chemical classes and modes of action were purchased from Sigma Chemicals (St. Louis, MO). Stock solutions were prepared by dissolving each antibiotic in an appropriate aqueous solution based on information provided by the manufacturer. The final working solution of each antibiotic contained no more than 0.2% (w/v) of any organic solvent. To determine their potency against a bacterial strain engineered for expression of an antisense complementary to a proliferation-required gene encoding 50S ribosomal protein, each antibiotic was serially diluted two or three fold in growth medium supplemented with the appropriate antibiotic for maintenance of the anti-sense construct. At least ten dilutions were prepared for each antibiotic. 25 μL aliquots of each dilution were transferred to discrete wells of a 384-well microplate (the assay plate) using a multi-channel pipette. Quadruplicate wells were used for each dilution of an antibiotic under each treatment condition (plus and minus inducer). Each assay plate contained twenty wells for cell growth controls (growth media replacing antibiotic), ten wells for each treatment (plus and minus inducer, in this example IPTG). Assay plates were usually divided into the two treatments: half the plate containing induced cells and an appropriate concentrations of inducer (in this example IPTG) to maintain the state of induction, the other half containing non-induced cells in the absence of IPTG.

Cells for the assay were prepared as follows. Bacterial cells containing a construct, from which expression of antisense nucleic acid complementary to *rplL* and *rplJ*, which encode proliferation-required 50S ribosomal subunit proteins, is inducible in the presence of IPTG, were grown into exponential growth (OD_{600} 0.2 to 0.3) and then diluted 1:100 into fresh media containing either 400 μM or 0 μM inducer (IPTG). These cultures were incubated at 37° C for 2.5 hr. After a 2.5 hr incubation, induced and non-induced cells were respectively diluted into an assay medium at a final OD_{600} value of 0.0004. The medium contained an appropriate concentration of the antibiotic for the maintenance of the anti-sense construct. In addition, the medium used to dilute induced cells was supplemented with 800 μM IPTG so that addition to the assay plate would result in a final IPTG concentration of 400 μM. Induced and non-induced cell suspensions were dispensed (25 μl/well) into the appropriate wells of the assay plate as

discussed previously. The plate was then loaded into a plate reader, incubated at constant temperature, and cell growth was monitored in each well by the measurement of light scattering at 595 nm. Growth was monitored every 5 minutes until the cell culture attained a stationary growth phase. For each concentration of antibiotic, a percentage inhibition of growth was calculated at the time point corresponding to mid-exponential growth for the associated control wells (no antibiotic, plus or minus IPTG). For each antibiotic and condition (plus or minus IPTG), a plot of percent inhibition versus log of antibiotic concentration was generated and the IC_{50} determined. A comparison of the IC_{50} for each antibiotic in the presence and absence of IPTG revealed whether induction of the antisense construct sensitized the cell to the mechanism of action exhibited by the antibiotic. Cells which exhibited a significant (standard statistical analysis) numerical decrease in the IC_{50} value in the presence of inducer were considered to have an increased sensitivity to the test antibiotic.

The results are provided in the table below, which lists the classes and names of the antibiotics used in the analysis, the targets of the antibiotics, the IC_{50} in the absence of IPTG, the IC_{50} in the presence of IPTG, the concentration units for the IC_{50}s, the fold increase in IC_{50} in the presence of IPTG, and whether increased sensitivity was observed in the presence of IPTG.

TABLE IV

Effect of Expression of Antisense RNA to rplL and rplJ on Antibiotic Sensitivity

ANTIBIOTIC CLASS /Names	TARGET	IC50 (-IPTG)	IC50 (+IPTG)	Conc. Unit	Fold Increase in Sensitivity	Sensitivity Increased?
PROTEIN SYNTHESIS INHIBITOR ANTIBIOTICS						
AMINOGLYCOSIDES						
Gentamicin	30S ribosome function	2715	19.19	ng/ml	141	Yes
Streptomycin	30S ribosome function	11280	161	ng/ml	70	Yes
Spectinomycin	30S ribosome function	18050	<156	ng/ml		Yes
Tobramycin	30S ribosome function	3594	70.58	ng/ml	51	Yes
MACROLIDES						
Erythromycin	50S ribosome function	7467	187	ng/ml	40	Yes
AROMATIC POYKETIDES						
Tetracycline	30S ribosome function	199.7	1.83	ng/ml	109	Yes
Minocycline	30S ribosome function	668.4	3.897	ng/ml	172	Yes
Doxycycline	30S ribosome function	413.1	27.81	ng/ml	15	Yes
OTHER PROTEIN SYNTHESIS INHIBITORS						
Fusidic acid	Elongation Factor G function	59990	641	ng/ml	94	Yes
Chloramphenicol	30S ribosome function	465.4	1.516	ng/ml	307	Yes
Lincomycin	50S ribosome function	47150	324.2	ng/ml	145	Yes
OTHER ANTIBIOTIC MECHANISMS						
B-LACTAMS						
Cefoxitin	Cell wall biosynthesis	2782	2484	ng/ml	1	No
Cefotaxime	Cell wall biosynthesis	24.3	24.16	ng/ml	1	No
DNA SYNTHESIS INHIBITORS						
Nalidixic acid	DNA Gyrase activity	6973	6025	ng/ml	1	No
Ofloxacin	DNA Gyrase activity	49.61	45.89	ng/ml	1	No
OTHER						
Bacitracin	Cell membrane function	4077	4677	mg/ml	1	No
Trimethoprim	Dihydrofolate Reductase activity	128.9	181.97	ng/ml	1	No
Vancomycin	Cell wall biosynthesis	145400	72550	ng/ml	2	No

The above results demonstrate that induction of an antisense RNA to genes encoding 50S ribosomal subunit proteins results in a selective and highly significant sensitization of cells to antibiotics that inhibit ribosomal function and protein synthesis. The above results further demonstrate that induction of an antisense construct to an essential gene sensitizes an organism to compounds that interfere with that gene products' biological role. This sensitization is restricted to compounds that interfere with pathways associated with the targeted gene and it's product.

Assays utilizing antisense constructs to essential genes can be used to identify compounds that specifically interfere with the activity of multiple targets in a pathway. Such constructs can be used to simultaneously screen a sample against multiple targets in one pathway in one reaction (Combinatorial HTS).

Furthermore, as discussed above, panels of antisense construct containing cells may be used to characterize the point of intervention of any compound affecting an essential biological pathway including antibiotics with no known mechanism of action.

Another embodiment of the present invention is a method for determining the pathway against which a test antibiotic compound is active in which the activity of target proteins or nucleic acids involved in proliferation-required pathways is reduced by contacting cells with a sub-lethal concentration of a known antibiotic which acts against the target protein or nucleic acid. In one embodiment, the target protein or nucleic acid is a target protein or nucleic acid corresponding to a proliferation-required nucleic acid identified using the methods described above. The method is similar to those described above for determining which pathway a test antibiotic acts against except that rather than reducing the activity or level of a proliferation-required gene product using a sub-lethal level of antisense to a proliferation-required nucleic acid, the activity or level of the proliferation-required gene product is reduced using sub-lethal level of a known antibiotic which acts against the proliferation required gene product.

Interactions between drugs which affect the same biological pathway has been described in the literature. For example, Mecillinam (Amdinocillin) binds to and inactivates the penicillin binding protein 2 (PBP2, product of the *mrdA* in *E. coli*). This antibiotic inteacts with other antibiotics that inhibit PBP2 as well as antibiotics that inhibit other penicillin binding proteins such as PBP3 [(Gutmann, L., Vincent, S., Billot-Klein, D., Acar, J.F., Mrena, E., and Williamson, R. (1986) Involvement of

penicillin-binding protein 2 with other penicillin-binding proteins in lysis of *Escherichia coli* by some beta-lactam antibiotics alone and in synergistic lytic effect of amdinocillin (mecillinam). Antimicrobial Agents & Chemotherapy, 30:906-912), the disclosure of which is incorporated herein by reference in its entirety]. Interactions between drugs could, therefore, involve two drugs that inhibit the same target protein or nucleic acid or inhibit different proteins or nucleic acids in the same pathway [(Fukuoka, T., Domon, H., Kakuta, M., Ishii, C., Hirasawa, A., Utsui, Y., Ohya, S., and Yasuda, H. (1997) Combination effect between panipenem and vancomycin on highly methicillin-resistant Staphylococcus aureus. Japan. J. Antibio. 50:411-419; Smith, C.E., Foleno, B.E., Barrett, J.F., and Frosc, M.B. (1997) Assessment of the synergistic interactions of levofloxacin and ampicillin against *Enterococcus faecium* by the checkerboard agar dilution and time-kill methods. Diagnos. Microbiol. Infect. Disease 27:85-92; den Hollander, J.G., Horrevorts, A.M., van Goor, M.L., Verbrugh, H.A., and Mouton, J.W. (1997) Synergism between tobramycin and ceftazidime against a resistant Pseudomonas aeruginosa strain, tested in an in vitro pharmacokinetic model. Antimicrobial Agents & Chemotherapy. 41:95-110), the disclosure of all of which are incorporated herein by reference in their entireties].

Two drugs may interact even though they inhibit different targets. For example, the proton pump inhibitor, Omeprazole, and the antibiotic, Amoxycillin, two synergistic compounds acting together, can cure *Helicobacter pylori* infection [(Gabryelewicz, A., Laszewicz, W., Dzieniszewski, J., Ciok, J., Marlicz, K., Bielecki, D., Popiela, T., Legutko, J., Knapik, Z., Poniewierka, E. (1997) Multicenter evaluation of dual-therapy (omeprazol and amoxycillin) for *Helicobacter pylori*-associated duodenal and gastric ulcer (two years of the observation). J. Physiol. Pharmacol. 48 Suppl 4:93-105), the disclosure of which is incorporated herein by reference in its entirety].

The growth inhibition from the sub-lethal concentration of the known antibiotic may be at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, or at least about 75%, or more.

Alternatively, the sub-lethal concentration of the known antibiotic may be determined by measuring the activity of the target proliferation-required gene product rather than by measuring growth inhibition.

Cells are contacted with a combination of each member of a panel of known antibiotics at a sub-lethal level and varying concentrations of the test antibiotic. As a control, the cells are contacted with varying concentrations of the test antibiotic alone. The IC_{50} of the test antibiotic in the presence and absence of the known antibiotic is determined. If the IC_{50}s in the presence and absence of the known drug are substantially similar, then the test drug and the known drug act on different pathways. If the IC_{50}s are substantially different, then the test drug and the known drug act on the same pathway.

Another embodiment of the present invention is a method for identifying a candidate compound for use as an antibiotic in which the activity of target proteins or nucleic acids involved in proliferation-required pathways is reduced by contacting cells with a sub-lethal concentration of a known antibiotic which acts against the target protein or nucleic acid. In one embodiment, the target protein or nucleic acid is a target protein or nucleic acid corresponding to a proliferation-required nucleic acid identified using the methods described above. The method is similar to those described above for identifying candidate compounds for use as antibiotics except that rather than reducing the activity or level of a proliferation-required gene product using a sub-lethal level of antisense to a proliferation-required nucleic acid, the activity or level of the proliferation-required gene product is reduced using a sub-lethal level of a known antibiotic which acts against the proliferation required gene product.

The growth inhibition from the sub-lethal concentration of the known antibiotic may be at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, or at least about 75%, or more.

Alternatively, the sub-lethal concentration of the known antibiotic may be determined by measuring the activity of the target proliferation-required gene product rather than by measuring growth inhibition.

In order to characterize test compounds of interest, cells are contacted with a panel of known antibiotics at a sub-lethal level and one or more concentrations of the test compound. As a control, the cells are contacted with the same concentrations of the test compound alone. The IC_{50} of the test compound in the presence and absence of the known antibiotic is determined. If the IC_{50} of the test compound is substantially

different in the presence and absence of the known drug then the test compound is a good candidate for use as an antibiotic. As discussed above, once a candidate compound is identified using the above methods its structure may be optimized using standard techniques such as combinatorial chemistry.

Representative known antibiotics which may be used in each of the above methods are provided in the table below. However, it will be appreciated that other antibiotics may also be used.

ANTIBIOTIC	INHIBITS/TARGET	RESISTANT MUTANTS
Inhibitors of Transcription		
Rifamycin, 1959 Rifampicin Rifabutin Rifaximin	Inhibits initiation of transcription/ß-subunit RNA polymerase, *rpoB*	*rpoB, crp, cyaA*
Streptolydigin	Accelerates transcription chain termination/ß-subunit RNA polymerase	*rpoB*
Streptovaricin	an acyclic ansamycin, inhibits RNA polymerase	*rpoB*
Actinomycin D+EDTA	Intercalates between 2 successive G-C pairs, *rpoB,* inhibits RNA synthesis	*pldA*
Inhibitors of Nucleic Acid Metabolism		
Quinolones, 1962 Nalidixic acid Oxolinic acid	α subunit gyrase and/or topoisomerase IV, *gyrA*	*gyrAorB, icd, sloB*
Fluoroquinolones Ciprofloxacin, 1983 Norfloxacin	α subunit gyrase, *gyrA* and/or topoisomerase IV (probable target in Staph)	*gyrA* *norA* (efflux in Staph) *hipQ*
Coumerins Novobiocin	Inhibits ATPase activity of ß-subunit gyrase, *gyrB*	*gyrB, cysB, cysE, nov, ompA*
Coumermycin	Inhibits ATPase activity of ß-subunit gyrase, *gyrB*	*gyrB, hisW*
Albicidin	DNA synthesis	*tsx* (nucleoside channel)
Metronidazole	Causes single-strand breaks in DNA	*nar*
Inhibitors of Metabolic Pathways		
Sulfonamides, 1932 Sulfanilamide	blocks synthesis of dihydrofolate,dihydro-pteroate synthesis, *folP*	*folP, gpt, pabA, pabB, pabC*
Trimethoprim, 1962	Inhibits dihydrofolate reductase, *folA*	*folA, thyA*
Showdomycin	Nucleoside analogue capable of alkylating sulfhydryl groups, inhibitor of thymidylate synthetase	*nupC, pnp*
Thiolactomycin	type II fatty acid synthase inhibitor	*emrB* *fadB, emrB* due to gene dosage
Psicofuranine	Adenosine glycoside antibiotic, target is GMP synthetase	*guaA,B*
Triclosan	Inhibits fatty acid synthesis	*fabI (envM)*

Diazoborines Isoniazid Ethionamide	heterocyclic, contains boron, inhibit fatty acid synthesis, enoyl-ACP reductase, *fabI*	*fabI (envM)*
Inhibitors of Translation		
Phenylpropanoids Chloramphenicol, 1947	Binds to ribosomal peptidyl transfer center preventing peptide translocation/ binds to S6, L3, L6, L14, L16, L25, L26, L27, but preferentially to L16	*rrn, cmlA, marA, ompF, ompR*
Tetracyclines, 1948, type II polyketides Minocycline Doxycycline	Binding to 30S ribosomal subunit, "A" site on 30S subunit, blocks peptide elongation, strongest binding to S7	*clmA (cmr), mar, ompF*
Macrolides (type I polyketides) Erythromycin, 1950 Carbomycin Spiramycin, etc	Binding to 50 S ribosomal subunit, 23S rRNA, blocks peptide translocation, L15, L4, L12	*rrn, rplC, rplD, rplV , mac*
Aminoglycosides Streptomycin, 1944 Neomycin Spectinomycin Kanamycin Kasugamycin Gentamicin, 1963 Amikacin Paromycin	Irreversible binding to 30S ribosomal subunit, prevents translation or causes mistranslation of mRNA/16S rRNA	*rpsL, strC,M, ubiF* *atpA-E, ecfB, hemAC,D,E,G, topA,* *rpsC,D,E, rrn, spcB* *atpA-atpE, cpxA, ecfB, hemA,B,L, topA* *ksgA,B,C,D, rplB,K, rpsI,N,M,R* *rplF, ubiF* *cpxA* *rpsL*
Lincosamides Lincomycin, 1955 Clindamycin	Binding to 50 S ribosomal subunit, blocks peptide translocation	*linB, rplN,O, rpsG*
Streptogramins Virginiamycin, 1955 Pristinamycin Synercid: quinupristin /dalfopristin	2 components, Streptogramins A&B, bind to the 50S ribosomal subunit blocking peptide translocation and peptide bond formation	
Fusidanes Fusidic Acid	Inhibition of elongation factor G (EF-G) prevents peptide translocation	*fusA*
Kirromycin (Mocimycin)	Inhibition of elongation factor TU (EF-Tu), prevents peptide bond formation	*tufA,B*
Pulvomycin	Binds to and inhibits EF-TU	
Thiopeptin	Sulfur-containing antibiotic, inhibits protein synthesis,EF-G	*rplE*
Tiamulin	Inhibits protein synthesis	*rplC, rplD*
Negamycin	Inhibits termination process of protein synthesis	*prfB*
Oxazolidinones Linezolid Isoniazid	23S rRNA	*pdx*
Nitrofurantoin	Inhibits protein synthesis, nitroreductases convert nitrofurantoin to highly reactive electrophilic intermediates which attack bacterial ribosomal proteins non-specifically	*nfnA,B*

Pseudomonic Acids Mupirocin (Bactroban)	Inhibition of isoleucyl tRNA synthetase- used for Staph, topical cream, nasal spray	*ileS*
Indolmycin	Inhibits tryptophanyl-tRNA synthetase	*trpS*
Viomycin		*rrmA* (23S rRNA methyltransferase; mutant has slow growth rate, slow chain elongation rate, and viomycin resistance)
Thiopeptides	Binds to L11-23S RNA complex	
Thiostrepton	Inhibits GTP hydrolysis by EF-G	
Micrococcin	Stimulates GTP hydrolysis by EF-G	

Inhibitors of Cell Walls/Membranes

ß-lactams	Inhibition of one or more cell wall	
Penicillin, 1929	transpeptidases, endopeptidases, and	
Ampicillin	glycosidases (PBPs), of the 12 PBPs only 2	*ampC, ampD, ampE,*
Methicillin, 1960	are essential: *mrdA* (PBP2) and *ftsI (pbpB,* PBP3)	*envZ, galU, hipA, hipQ, ompC, ompF, ompR, ptsI, rfa, tolD, tolE*
Cephalosporins, 1962		*tonB*
Mecillinam (amdinocillin)	Binds to and inactivates PBP2 (*mrdA*)	*alaS, argS, crp, cyaA, envB, mrdA,B, mreB,C,D*
Aztreonam (Furazlocillin)	Inactivates PBP3 (*ftsI*)	
Bacilysin, Tetaine	Dipeptide, inhib glucosamine synthase	*dppA*
Glycopeptides	Inhib G+ cell wall syn, binds to terminal	
Vancomycin, 1955	D-ala-D-ala of pentapeptide,	
Polypeptides	Prevents dephosphorylation and	
Bacitracin	regeneration of lipid carrier	*rfa*
Cyclic lipopeptide	Disrupts multiple aspects of membrane	
Daptomycin, 1980	function, including peptidoglycan synthesis, lipoteichoic acid synthesis, and the bacterial membrane potential	
Cyclic polypeptides	Surfactant action disrupts cell membrane	*pmrA*
Polymixin, 1939	lipids, binds lipid A mioety of LPS	
Fosfomycin, 1969	Analogue of P-enolpyruvate, inhibits 1st step in peptidoglycan synthesis - UDP-N-acetylglucosamine enolpyruvyl transferase, *murA*. Also acts as Immunosuppressant	*murA, crp, cyaA glpT, hipA, ptsI, uhpT*
Cycloserine	Prevents formation of D-ala dimer, inhibits D-ala ligase, *ddlA,B*	*hipA, cycA*
Alafosfalin	phosphonodipeptide, cell wall synthesis inhibitor, potentiator of β-lactams	*pepA, tpp*

Inhibitors of Protein Processing/Transport

Globomycin	Inhibits signal peptidase II (cleaves prolipoproteins subsequent to lipid modification, *lspA*	*lpp, dnaE*

EXAMPLE 12

Transfer of Exogenous Nucleic Acid Sequences to other Bacterial Species Using the *E. coli* Expression Vectors or Expression Vectors Functional in Bacterial Species other than *E. coli.*

Molecule No. EcXA190, encoding a portion of the *b3052* gene of *Escherichia coli*, was either transformed directly into *Enterobacter cloacae, Salmonella typhimurium* and/or *Klebsiella pneumoniae* or subcloned into an expression vector functional in these species and the subclones transformed into these organisms. Suitable expression vectors are well known in the art. These expression vectors were introduced into *Enterobacter cloacae, Salmonella typhimurium* and/or *Klebsiella pneumoniae* cells that were then assayed for growth inhibition according to the method of Example 1. After growth in liquid culture, cells were plated at various serial dilutions and a score determined by calculating the log difference in growth for INDUCED vs. UNINDUCED antisense RNA expression as determined by the maximum 10 fold dilution at which a colony was observed. If there was no effect of antisense RNA expression in one organism, the clone is given a score of zero "0" in that organism. In contrast, a score of "8" means that 10^8 times more cells were required to observe a colony formed on the induced state than in the non-induced state under the conditions used and in that organism.

Expression vectors containing Molecule No. EcXA190 were found to inhibit bacterial growth in all four organisms when expression of the antisense RNA was induced with IPTG. A score of 4 was assigned for *Escherichia coli,* 6 for *Enterobacter cloacae*, and 8 for *Salmonella typhimurium* and 3 for *Klebsiella pneumoniae* (obvious additional growth defect as well). The protein encoded by this sequence may be used as a target sequence to screen candidate compound libraries as described above.

In addition, the above methods were validated using other antisense nucleic acids which inhibit the growth of *E. coli* which were identified using methods similar to those described above. Expression vectors which inhibited growth of *E. coli* upon induction of antisense RNA expression with IPTG were transformed directly into *Enterobacter cloacae, Klebsiella pneumonia* or *Salmonella typhimurium.* The transformed cells were then assayed for growth inhibition according to the method of

Example 1. After growth in liquid culture, cells were plated at various serial dilutions and a score determined by calculating the log difference in growth for INDUCED vs. UNINDUCED antisense RNA expression as determined by the maximum 10 fold dilution at which a colony was observed. The results of these experiments are listed in Table V below. If there was no effect of antisense RNA expression in a microorganism, the clone is minus in Table V below. In contrast, a positive in Table V below means that at least 10 fold more cells were required to observe a colony on the induced plate than on the non-induced plate under the conditions used and in that microorganism.

TABLE V

Sensitivity of Other Microorganisms to Antisense Nucleic Acids That Inhibit Proliferation in *E. coli*

Mol. No.	*S. typhimurium*	*E. cloacae*	*K. pneumoniae*
EcXA001	+	+	-
EcXA004	+	-	-
EcXA005	+	+	+
EcXA006	-	-	-
EcXA007	-	+	-
EcXA008	+	-	+
EcXA009	-	-	-
EcXA010	+	+	+
EcXA011	-	+	-
EcXA012	-	+	-
EcXA013	+	+	+
EcXA014	+	+	-
EcXA015	+	+	+
EcXA016	+	+	+
EcXA017	+	+	+
EcXA018	+	+	+
EcXA019	+	+	+
EcXA020	+	+	+
EcXA021	+	+	+
EcXA023	+	+	+
EcXA024	+	-	+
EcXA025	-	-	-
EcXA026	+	+	-
EcXA027	+	+	-
EcXA028	+	-	-
EcXA029	-	-	-
EcXA030	+	+	+
EcXA031	+	-	-
EcXA032	+	+	-
EcXA033	+	+	+

EcXA034	+	+	+
EcXA035	-	-	-
EcXA036	+	-	+
EcXA037	+	+	-
EcXA038	+	+	+
EcXA039	+	-	-
EcXA041	+	+	+
EcXA042	-	+	+
EcXA043	-	-	-
EcXA044	-	-	-
EcXA045	+	+	+
EcXA046	-	-	-
EcXA047	+	+	-
EcXA048	-	-	-
EcXA049	+	-	-
EcXA050	-	-	-
EcXA051	+	-	-
EcXA052	+	-	-
EcXA053	+	+	+
EcXA054	-	-	+
EcXA055	+	-	-
EcXA056	+	-	+
EcXA057	+	+	-
EcXA058	-	-	-
EcXA059	+	+	+
EcXA060	-	-	-
EcXA061	-	-	-
EcXA062	-	-	-
EcXA063	+	+	-
EcXA064	-	-	-
EcXA065	+	+	-
EcXA066	-	-	-
EcXA067	-	+	-
EcXA068	-	-	-
EcXA069	-	+	-
EcXA070	-	-	-
EcXA071	+	-	-
EcXA072	+	-	+
EcXA073	+	+	+
EcXA074	+	+	+
EcXA075	+	-	-
EcXA076	-	+	-
EcXA077	+	+	-
EcXA079	+	+	+
EcXA080	+	-	-
EcXA082	-	+	-

EcXA083	-	-	-
EcXA084	-	+	-
EcXA086	-	-	-
EcXA087	-	-	-
EcXA088	-	-	-
EcXA089	-	-	-
EcXA090	-	-	-
EcXA091	-	-	-
EcXA092	-	-	-
EcXA093	-	-	-
EcXA094	+	+	+
EcXA095	+	+	-
EcXA096	-	-	-
EcXA097	+	-	-
EcXA098	+	-	-
EcXA099	-	-	-
EcXA100	-	-	-
EcXA101	-	-	-
EcXA102	-	-	-
EcXA103	-	+	-
EcXA104	+	+	+
EcXA106	+	+	-
EcXA107	-	-	-
EcXA108	-	-	-
EcXA109	-	-	-
EcXA110	+	+	-
EcXA111	-	-	-
EcXA112	-	+	-
EcXA113	+	+	+
EcXA114	-	+	-
EcXA115	-	+	-
EcXA116	+	+	-
EcXA117	+	-	-
EcXA118	-	-	-
EcXA119	+	+	-
EcXA120	-	-	-
EcXA121	-	-	-
EcXA122	+	-	+
EcXA123	+	-	-
EcXA124	-	-	-
EcXA125	-	-	-
EcXA126	-	-	-
EcXA127	+	+	-
EcXA128	-	-	-
EcXA129	-	+	-
EcXA130	+	+	-
EcXA132	-	-	-

EcXA133	-	-	-
EcXA136	-	-	-
EcXA137	-	-	-
EcXA138	+	-	-
EcXA139	-	-	-
EcXA140	+	-	-
EcXA141	+	-	-
EcXA142	-	-	-
EcXA143	-	+	-
EcXA144	+	+	-
EcXA145	-	-	-
EcXA146	-	-	-
EcXA147	-	-	-
EcXA148	-	-	-
EcXA149	+	+	+
EcXA150	-	-	-
EcXA151	+	-	-
EcXA152	-	-	-
EcXA153	+	+	-
EcXA154	-	-	-
EcXA155	-	-	ND
EcXA156	-	+	-
EcXA157	-	-	-
EcXA158	-	-	-
EcXA159	+	-	-
EcXA160	+	-	-
EcXA162	-	-	-
EcXA163	-	-	-
EcXA164	-	-	-
EcXA165	-	-	-
EcXA166	-	-	-
EcXA167	-	-	-
EcXA168	-	-	-
EcXA169	-	+	-
EcXA171	-	-	-
EcXA172	-	-	-
EcXA173	-	-	-
EcXA174	-	-	-
EcXA175	-	-	-
EcXA176	-	-	-
EcXA178	-	-	-
EcXA179	-	-	-
EcXA180	+	-	-
EcXA181	-	-	-
EcXA182	-	-	-
EcXA183	-	-	-
EcXA184	-	-	-

EcXA185	-	-	-
EcXA186	-	-	-
EcXA187	+	+	+
EcXA189	+	-	-
EcXA190	+	+	+
EcXA191	+	+	-
EcXA192	-	+	-

Thus, the ability of an antisense nucleic acid which inhibits the proliferation of *E. coli* to inhibit the growth of other organims may be evaluated by either transforming the antisense nucleic acid directly into other *Escherichia* species or inserting the antisense nucleic acid into expression vectors that are functional in other Gram negative species such as *Enterobacter cloacae, Salmonella typhimurium, and/or Klebsiella pneumoniae* . Similarly, the antisense nucleic acid can be inserted in expression vectors that are functional in Gram-positive species such as *Staphylococcus aureus, Enterococcus faecalis* and *Streptococcus pneumoniae* or other species.

Those skilled in the art will appreciate that a negative result in a heterologous microorganism does not mean that that microorganism is missing that gene nor does it mean that the gene is unessential. However, a positive result means that the heterologous microorganism contains a homologous gene which is required for proliferation of that microorganism. The homologous gene may be obtained using the methods described herein. Those cells that are inhibited by antisense may be used in cell-based assays as described herein for the identification and characterization of compounds in order to develop antibiotics effective in these microorganisms. Those skilled in the art will appreciate that an antisense molecule which works in the microorganism from which it was obtained will not always work in a heterologous microorganism.

EXAMPLE 13

Use of Identified Exogenous Nucleic Acid Sequences as Probes

The identified sequence of the present invention can be used as probes to obtain the sequence of additional genes of interest from a second organism. For example, probes to genes encoding potential bacterial target proteins may be hybridized to nucleic acids from other organisms including other bacteria and higher organisms, to identify homologous sequences. Such hybridization might indicate that the protein encoded by the gene to which the probe corresponds is found in humans and therefore not necessarily a good drug target. Alternatively, the gene can be conserved only in

bacteria and therefore would be a good drug target for a broad spectrum antibiotic or antimicrobial.

Probes derived from the identified nucleic acid sequences of interest or portions thereof can be labeled with detectable labels familiar to those skilled in the art, including radioisotopes and non-radioactive labels, to provide a detectable probe. The detectable probe can be single stranded or double stranded and can be made using techniques known in the art, including *in vitro* transcription, nick translation, or kinase reactions. A nucleic acid sample containing a sequence capable of hybridizing to the labeled probe is contacted with the labeled probe. If the nucleic acid in the sample is double stranded, it can be denatured prior to contacting the probe. In some applications, the nucleic acid sample can be immobilized on a surface such as a nitrocellulose or nylon membrane. The nucleic acid sample can comprise nucleic acids obtained from a variety of sources, including genomic DNA, cDNA libraries, RNA, or tissue samples.

Procedures used to detect the presence of nucleic acids capable of hybridizing to the detectable probe include well known techniques such as Southern blotting, Northern blotting, dot blotting, colony hybridization, and plaque hybridization. In some applications, the nucleic acid capable of hybridizing to the labeled probe can be cloned into vectors such as expression vectors, sequencing vectors, or in vitro transcription vectors to facilitate the characterization and expression of the hybridizing nucleic acids in the sample. For example, such techniques can be used to isolate, purify and clone sequences from a genomic library, made from a variety of bacterial species, which are capable of hybridizing to probes made from the sequences identified in Examples 5 and 6.

EXAMPLE 14

Preparation of PCR Primers and Amplification of DNA

The identified *E. coli* genes corresponding directly to or located within the operon of nucleic acid sequences required for proliferation or portions thereof can be used to prepare PCR primers for a variety of applications, including the identification or isolation of homologous sequences from other species, for example *S. typhimurium, E. cloacae, E. faecalis, S. pneumoniae, and K. pneumoniae*, which contain part or all of the homologous genes. Because homologous genes are related but not identical in sequence, those skilled in the art will often employ degenerate sequence PCR primers. Such degenerate sequence primers are designed based on conserved sequence regions, either known or suspected,

such as conserved coding regions. The successful production of a PCR product using degenerate probes generated from the sequences identified herein would indicate the presence of a homologous gene sequence in the species being screened. The PCR primers are at least 10 nucleotides, and preferably at least 20 nucleotides in length. More preferably, the PCR primers are at least 20-30 nucleotides in length. In some embodiments, the PCR primers can be more than 30 nucleotides in length. It is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approximately the same. A variety of PCR techniques are familiar to those skilled in the art. For a review of PCR technology, see Molecular Cloning to Genetic Engineering White, B.A. Ed. in Methods in Molecular Biology 67: Humana Press, Totowa 1997. When the entire coding sequence of the target gene is known, the 5' and 3' regions of the target gene can be used as the sequence source for PCR probe generation. In each of these PCR procedures, PCR primers on either side of the nucleic acid sequences to be amplified are added to a suitably prepared nucleic acid sample along with dNTPs and a thermostable polymerase such as Taq polymerase, Pfu polymerase, or Vent polymerase. The nucleic acid in the sample is denatured and the PCR primers are specifically hybridized to complementary nucleic acid sequences in the sample. The hybridized primers are extended. Thereafter, another cycle of denaturation, hybridization, and extension is initiated. The cycles are repeated multiple times to produce an amplified fragment containing the nucleic acid sequence between the primer sites.

EXAMPLE 15

Inverse PCR

The technique of inverse polymerase chain reaction can be used to extend the known nucleic acid sequence identified in Examples 5 and 6. The inverse PCR reaction is described generally by Ochman et al., in Ch. 10 of PCR Technology: Principles and Applications for DNA Amplification, (Henry A. Erlich, Ed.) W.H. Freeman and Co. (1992). Traditional PCR requires two primers that are used to prime the synthesis of complementary strands of DNA. In inverse PCR, only a core sequence need be known.

Using the sequences identified as relevant from the techniques taught in Examples 5 and 6 and applied to other species of bacteria, a subset of exogenous nucleic sequences are identified that correspond to genes or operons that are required for bacterial proliferation. In species for which a genome sequence is not known, the technique of

inverse PCR provides a method for obtaining the gene in order to determine the sequence or to place the probe sequences in full context to the target sequence to which the identified exogenous nucleic acid sequence binds.

To practice this technique, the genome of the target organism is digested with an appropriate restriction enzyme so as to create fragments of nucleic acid that contain the identified sequence as well as unknown sequences that flank the identified sequence. These fragments are then circularized and become the template for the PCR reaction. PCR primers are designed in accordance with the teachings of Example 15 and directed to the ends of the identified sequence. The primers direct nucleic acid synthesis away from the known sequence and toward the unknown sequence contained within the circularized template. After the PCR reaction is complete, the resulting PCR products can be sequenced so as to extend the sequence of the identified gene past the core sequence of the identified exogenous nucleic acid sequence identified. This process can be repeated iteratively if necessary. In this manner, the full sequence of each novel gene can be identified. Additionally the sequences of adjacent coding and noncoding regions can be identified.

EXAMPLE 16

Identification of Genes Required for *Staphylococcus aureus* Proliferation

Genes required for proliferation in *Staphylococcus aureus* are identified according to the methods described above.

EXAMPLE 17

Identification of Genes Required for *Neisseria gonorrhoeae* Proliferation

Genes required for proliferation in *Neisseria gonorrhoeae* are identified according to the methods described above.

EXAMPLE 18

Identification of Genes Required for *Pseudomonas aeruginosa* Proliferation

Genes required for proliferation in *Pseudomonas aeruginosa* are identified according to the methods described above.

EXAMPLE 19

Identification of Genes Required for *Enterococcus faecalis* Proliferation

Genes required for proliferation in *Enterococcus faecalis* are identified according to the methods described above.

EXAMPLE 20

Identification of Genes Required for *Haemophilus influenzae* Proliferation

Genes required for proliferation in *Haemophilus influenzae* are identified according to the methods described above.

EXAMPLE 21

Identification of Genes Required for *Salmonella typhimurium* Proliferation

Genes required for proliferation in *Salmonella typhimurium* are identified according to the methods described above.

EXAMPLE 22

Identification of Genes Required for *Helicobacter pylori* Proliferation

Genes required for proliferation in *Helicobacter pylori* are identified according to the methods described above.

EXAMPLE 23

Identification of Genes Required for *Mycoplasma pneumoniae* Proliferation

Genes required for proliferation in *Mycoplasma pneumoniae* are identified according to the methods described above.

EXAMPLE 24

Identification of Genes Required for *Plasmodium ovale* Proliferation

Genes required for proliferation in *Plasmodium ovale* are identified according to the methods described above.

EXAMPLE 25

Identification of Genes Required for *Saccharomyces cerevisiae* Proliferation

Genes required for proliferation in *Saccharomyces cerevisiae* are identified according to the methods described above.

EXAMPLE 26

Identification of Genes Required for *Entamoeba histolytica* Proliferation

Genes required for proliferation in *Entamoeba histolytica* are identified according to the methods described above.

EXAMPLE 27

Identification of Genes Required for *Candida albicans* Proliferation

Genes required for proliferation in *Candida albicans* are identified according to the methods described above.

EXAMPLE 28

Identification of Genes Required for *Klebsiella pneumoniae* Proliferation

Genes required for proliferation in *Klebsiella pneumoniae* are identified according to the methods described above.

EXAMPLE 29

Identification of Genes Required for *Salmonella typhi* Proliferation

Genes required for proliferation in *Salmonella typhi* are identified according to the methods described above.

EXAMPLE 30

Identification of Genes Required for *Salmonella paratyphi* Proliferation

Genes required for proliferation in *Salmonella paratyphi* are identified according to the methods described above.

EXAMPLE 31

Identification of Genes Required for *Salmonella cholerasuis* Proliferation

Genes required for proliferation in *Salmonella cholerasuis* are identified according to the methods described above.

EXAMPLE 32

Identification of Genes Required for *Staphylococcus epidermis* Proliferation

Genes required for proliferation in *Staphylococcus epidermis* are identified according to the methods described above.

EXAMPLE 33

Identification of Genes Required for *Mycobacterium tuberculosis* Proliferation

Genes required for proliferation in *Mycobacterium tuberculosis* are identified according to the methods described above.

EXAMPLE 34

Identification of Genes Required for *Mycobacterium leprae* Proliferation

Genes required for proliferation in *Mycobacterium leprae* are identified according to the methods described above.

EXAMPLE 35

Identification of Genes Required for *Treponema pallidum* Proliferation

Genes required for proliferation in *Treponema pallidum* are identified according to the methods described above.

EXAMPLE 36

Identification of Genes Required for *Bacillus anthracis* Proliferation

Genes required for proliferation in *Bacillus anthracis* are identified according to the methods described above.

EXAMPLE 37

Identification of Genes Required for *Yersinia pestis* Proliferation

Genes required for proliferation in *Yersinia pestis* are identified according to the methods described above.

EXAMPLE 38

Identification of Genes Required for *Clostridium botulinum* Proliferation

Genes required for proliferation in *Clostridium botulinum* are identified according to the methods described above.

EXAMPLE 39

Identification of Genes Required for *Campylobacter jejuni* Proliferation

Genes required for proliferation in *Campylobacter jejuni* are identified according to the methods described above.

EXAMPLE 40

Identification of Genes Required for *Chlamydia trachomatis* Proliferation

Genes required for proliferation in *Chlamydia trachomatis* are identified according to the methods described above.

It will be appreciated that genes required for proliferation of any microorganism of interest, including those specifically mentioned herein, may be identified according to the methods described above.

Use of Isolated Exogenous Nucleic Acid Fragments as Antisense Antibiotics

In addition to using the identified sequences to enable screening of molecule libraries to identify compounds useful to identify antibiotics, the sequences themselves can be used as therapeutic agents. Specifically, the identified exogenous sequences in an antisense orientation can be provided to an individual to inhibit the translation of a bacterial target gene.

Generation of Antisense Therapeutics from Identified Exogenous Sequences

The sequences of the present invention can be used as antisense therapeutics for the treatment of bacterial infections or simply for inhibition of bacterial growth *in vitro* or *in vivo*. The therapy exploits the biological process in cells where genes are transcribed into messenger RNA (mRNA) that is then translated into proteins. Antisense RNA technology contemplates the use of antisense oligonucleotides complementary to a target gene that will bind to its target nucleic acid and decrease or inhibit the expression of the target gene. For example, the antisense nucleic acid may inhibit the translation or transcription of the target nucleic acid. In one embodiment, antisense oligonucleotides can be used to treat and control a bacterial infection of a cell culture containing a population of desired cells contaminated with bacteria. In another embodiment, the antisense oligonucleotides can be used to treat an organism with a bacterial infection.

Antisense oligonucleotides can be synthesized from any of the sequences of the present invention using methods well known in the art. In a preferred embodiment, antisense oligonucleotides are synthesized using artificial means. Uhlmann & Peymann, Chemical Rev. 90:543-584 (1990) review antisense oligonucleotide technology in detail. Modified or unmodified antisense oligonucleotides can be used as therapeutic agents. Modified antisense oligonucleotides are preferred since it is well known that antisense oligonucleotides are extremely unstable. Modification of the phosphate backbones of the antisense oligonucleotides can be achieved by substituting the internucleotide phosphate residues with methylphosphonates, phosphorothioates, phosphoramidates, and phosphate esters. Nonphosphate internucleotide analogs such as siloxane bridges, carbonate brides, thioester bridges, as well as many others known in the art may also be used. The preparation of certain antisense oligonucleotides with modified internucleotide linkages is described in U.S. Patent No. 5,142,047, hereby incorporated by reference.

Modifications to the nucleoside units of the antisense oligonucleotides are also contemplated. These modifications can increase the half-life and increase cellular rates of uptake for the oligonucleotides *in vivo*. For example, α-anomeric nucleotide units and modified nucleotides such as 1,2-dideoxy-d-ribofuranose, 1,2-dideoxy-1-

phenylribofuranose, and N^4, N^4-ethano-5-methyl-cytosine are contemplated for use in the present invention.

An additional form of modified antisense molecules is found in peptide nucleic acids. Peptide nucleic acids (PNA) have been developed to hybridize to single and double stranded nucleic acids. PNA are nucleic acid analogs in which the entire deoxyribose-phosphate backbone has been exchanged with a chemically completely different, but structurally homologous, polyamide (peptide) backbone containing 2-aminoethyl glycine units. Unlike DNA, which is highly negatively charged, the PNA backbone is neutral. Therefore, there is much less repulsive energy between complementary strands in a PNA-DNA hybrid than in the comparable DNA-DNA hybrid, and consequently they are much more stable. PNA can hybridize to DNA in either a Watson/Crick or Hoogsteen fashion (Demidov et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**:2637-2641, 1995; Egholm, *Nature* **365**:566-568, 1993; Nielsen et al., *Science* **254**:1497-1500, 1991; Dueholm et al., *New J. Chem.* **21**:19-31, 1997).

Molecules called PNA "clamps" have been synthesized which have two identical PNA sequences joined by a flexible hairpin linker containing three 8-amino-3,6-dioxaoctanoic acid units. When a PNA clamp is mixed with a complementary homopurine or homopyrimidine DNA target sequence, a PNA-DNA-PNA triplex hybrid can form which has been shown to be extremely stable (Bentin et al., *Biochemistry* **35**:8863-8869, 1996; Egholm et al., *Nucleic Acids Res.* **23**:217-222, 1995; Griffith et al., *J. Am. Chem. Soc.* **117**:831-832, 1995).

The sequence-specific and high affinity duplex and triplex binding of PNA have been extensively described (Nielsen et al., *Science* **254**:1497-1500, 1991; Egholm et al., *J. Am. Chem. Soc.* **114**:9677-9678, 1992; Egholm et al., *Nature* **365**:566-568, 1993; Almarsson et al., *Proc. Natl. Acad. Sci. U.S.A.* **90**:9542-9546, 1993; Demidov et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**:2637-2641, 1995). They have also been shown to be resistant to nuclease and protease digestion (Demidov et al., *Biochem. Pharm.* **48**:1010-1313, 1994). PNA has been used to inhibit gene expression (Hanvey et al., *Science* **258**:1481-1485,1992; Nielsen et al., *Nucl. Acids. Res.*, **21**:197-200, 1993; Nielsen et al., *Gene* **149**:139-145, 1994; Good & Nielsen, **Science**, 95: 2073-2076, 1998; all of which are hereby incorporated by reference), to block restriction enzyme activity (Nielsen et al., *supra.*, 1993), to act as an artificial transcription promoter (Mollegaard, *Proc. Natl. Acad.*

Sci. U.S.A. **91**:3892-3895, 1994) and as a pseudo restriction endonuclease (Demidov et al., *Nucl. Acids. Res.* **21**:2103-2107, 1993). Recently, PNA has also been shown to have antiviral and antitumoral activity mediated through an antisense mechanism (Norton, *Nature Biotechnol.*, **14**:615-619, 1996; Hirschman et al., *J. Investig. Med.* **44**:347-351, 1996). PNAs have been linked to various peptides in order to promote PNA entry into cells (Basu et al., *Bioconj. Chem.* **8**:481-488, 1997; Pardridge et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**:5592-5596, 1995).

The antisense oligonucleotides contemplated by the present invention can be administered by direct application of oligonucleotides to a target using standard techniques well known in the art. The antisense oligonucleotides can be generated within the target using a plasmid, or a phage. Alternatively, the antisense nucleic acid may be expressed from a sequence in the chromosome of the target cell. For example, a promoter may be introduced into the chromosme of the target cell near the target gene such that the promoter directs teh transcription of the antisense nucleic acid. Alternatively, a nucleic acid containing the antisense sequence operably linked to a promoter may be introduced into the chromosome of the target cell. It is further contemplated that contemplated that the antisense oligonucleotide contemplated are incorporated in a ribozyme sequence to enable the antisense to specifically bind and cleave its target mRNA. For technical applications of ribozyme and antisense oligonucleotides see Rossi et al., Pharmacol. Ther. 50(2):245-254, (1991), which is hereby incorporated by reference. The present invention also contemplates using a retron to introduce an antisense oligonucleotide to a cell. Retron technology is exemplified by U.S. Patent No. 5,405,775, which is hereby incorporated by reference. Antisense oligonucleotides can also be delivered using liposomes or by electroporation techniques which are well known in the art.

The antisense nucleic acids of the present invention can also be used to design antibiotic compounds comprising nucleic acids which function by intracellular triple helix formation. Triple helix oligonucleotides are used to inhibit transcription from a genome. The sequences identified as required for proliferation in the present invention, or portions thereof, can be used as templates to inhibit microorganism gene expression in individuals infected with such organisms. Traditionally, homopurine sequences were considered the most useful for triple helix strategies. However, homopyrimidine sequences can also

inhibit gene expression. Such homopyrimidine oligonucleotides bind to the major groove at homopurine:homopyrimidine sequences. Thus, both types of sequences based on the sequences of the present invention that are required for proliferation are contemplated for use as antibiotic compound templates.

The antisense oligonucleotides of this example employ the identified sequences of the present invention to induce bacterial cell death or at least bacterial stasis by inhibiting target nucleic acid transcription or translation. Antisense oligonucleotides containing from about 8 to 40 nucleotides of the sequences of the present invention have sufficient complementary to form a duplex with the target sequence under physiological conditions.

To kill bacterial cells or inhibit their growth, the antisense oligonucleotides are applied to the bacteria or to the target cells under conditions that facilitate their uptake. These conditions include sufficient incubation times of cells and oligonucleotides so that the antisense oligonucleotides are taken up by the cells. In one embodiment, an incubation period of 7-10 days is sufficient to kill bacteria in a sample. An optimum concentration of antisense oligonucleotides is selected for use.

The concentration of antisense oligonucleotides to be used can vary depending on the type of bacteria sought to be controlled, the nature of the antisense oligonucleotide to be used, and the relative toxicity of the antisense oligonucleotide to the desired cells in the treated culture. Antisense oligonucleotides can be introduced to cell samples at a number of different concentrations preferably between $1x10^{-10}$M to $1x10^{-4}$M. Once the minimum concentration that can adequately control gene expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo*. For example, an inhibiting concentration in culture of $1x10^{-7}$ translates into a dose of approximately 0.6 mg/kg body weight. Levels of oligonucleotide approaching 100 mg/kg body weight or higher may be possible after testing the toxicity of the oligonucleotide in laboratory animals. It is additionally contemplated that cells from the subject are removed, treated with the antisense oligonucleotide, and reintroduced into the subject. This range is merely illustrative and one of skill in the art are able to determine the optimal concentration to be used in a given case.

After the bacterial cells have been killed or controlled in a desired culture, the desired cell population may be used for other purposes.

EXAMPLE 41

The following example demonstrates the ability of an *E. coli* antisense oligonucleotide to act as a bactericidal or bacteriostatic agent to treat a contaminated cell culture system. The application of the antisense oligonucleotides of the present invention are thought to inhibit the translation of bacterial gene products required for proliferation.

The antisense oligonucleotide of this example corresponds to a 30 base phophorothioate modified oligodeoxynucelotide complementary to a nucleic acid involved in proliferation, such as Molecule Number EcXA118 (SEQ ID NO: 1). A sense oligodeoxynucelotide complementary to the antisense sequence is synthesized and used as a control. The oligonucleotides are synthesized and purified according to the procedures of Matsukura, et al., Gene 72:343 (1988). The test oligonucleotides are dissolved in a small volume of autoclaved water and added to culture medium to make a 100 micromolar stock solution.

Human bone marrow cells are obtained from the peripheral blood of two patients and cultured according standard procedures well known in the art. The culture is contaminated with the K-12 strain of *E. coli* and incubated at 37°C overnight to establish bacterial infection.

The control and antisense oligonucleotide containing solutions are added to the contaminated cultures and monitored for bacterial growth. After a 10 hour incubation of culture and oligonucleotides, samples from the control and experimental cultures are drawn and analyzed for the translation of the target bacterial gene using standard microbiological techniques well known in the art. The target *E. coli* gene is found to be translated in the control culture treated with the control oligonucleotide, however, translation of the target gene in the experimental culture treated with the antisense oligonucleotide of the present invention is not detected or reduced.

One way to determine if a gene is essential for proliferation in a host or virulence in a host is to construct a conditional allele of the gene an infectious organism. The host is then challenged with the organism under conditions in which the product of the gene is functional or non-functional or has reduced activity or where the gene product is absent or else present but at a reduced level. If the gene is essential for proliferation or virulence the infection of the host will be diminished or abolished under conditions in which the product

of the gene is not functional or has reduced activity or where the gene product is absent or else present but at a reduced level.

Since expression of an antisense nucleic acid complementary to a gene required for proliferation also decreases the synthesis of the gene product, antisense nucleic acids may also be used to evaluate whether a gene is essential for the proliferation or virulence of an infectious organism in the host. In such methods, nucleic acids encoding an antisense molecule complementary to the desired target gene are introduced into the infectious organism. For exampe, plasmids comprising one of SEQ ID NOs: 1-93 or fragments thereof which inhibit proliferation, may be introduced into the infectious organism. In some embodiments, the antisense nucleic acid may be transcribed from the IPTG-inducible promoter in pLEX5BA or from other regulated promoters or vector systems.

E. coli is transformed with the nucleic acid encoding the antisense molecule by electroporation and grown in medium which selects for the presence of the vector from which the antisense nucleic acid is expressed. The essentiality of the target for each antisense nucleic acid is verified in microorganisms grown in culture using the techniques described herein.

The ability of antisense expression to block *E. coli* infection in an animal is tested using the rabbit model of bacterial meningitis. A spinal needle is surgically placed into the cisterna magna of New Zealand White rabbits. The rabbits are inoculated with 10^5 to 10^6 cells of a normally virulent *E. coli* strain expressing an antisense nucleic acid complementary to a gene required for proliferation. Repeated CSF sampling is undertaken to determine multiple parameters of injury and infection such as cytochemical abnormalities, intracranial pressure, cerebral edema, BBB permeability, cerebral perfusion pressure and recovery of viable *E. coli* cells. Control animals are given intravenous injections of saline, which will not induce expression of the antisense nucleic acid, while experimental animals are given IPTG in intravenous injections to induce expression of the antisense nucleic acid. Alternatively, expression of the antisense nucleic acid may be induced by intravenous infusion of IPTG at sub-toxic levels. If other promoters other than IPTG inducible promoters are used, the rabbits may be fed the inducer in their water.

The use of rabbits allows multiple CSF samples per animal (one rabbit can give up to 8 sequential samples without change in CSF pressure). Treated animals receive therapy

from 2 hours post receiving inoculation up to several days. A typical efficacy study consists of 3 control animals and 3 treated animals.

The control animals in which expression of the antisense nucliec acid is not induced are not protected against infection with *E. coli* and there is a logarithmic increase in viable bacteria. In experimental animals, *E. coli* cells recovered from the site of infection are viable until antisense expression is subsequently induced. However, if the antisense nucleic acid is directed against a gene required for proliferation, after treatment with the inducer for antisense expression the *E. coli* cells infecting these rabbits will not multiply and fewer viable cells will be recovered from the site of infection. The *E. coli* cells recovered from the rabbits treated with the inducer are recovered, if still present, and assayed as above to determine if the promoter and gene are still present and functional. Conversely, if the antisense nucleic acid is not complementary to a gene required for proliferation, treatment of the rabbits with inducer will have no effect on *E. coli* viability.

EXAMPLE 42

A subject suffering from an *E. coli* infection is treated with the antisense oligonucleotide preparation of Example 39. The antisense oligonucleotide is provided in a pharmaceutically acceptable carrier at a concentration effective to inhibit the transcription or translation of the target nucleic acid. The present subject is treated with a concentration of antisense oligonucleotide sufficient to achieve a blood concentration of about 100 micromolar. The patient receives daily injections of antisense oligonucleotide to maintain this concentration for a period of 1 week. At the end of the week a blood sample is drawn and analyzed for the presence or absence of the orgranism using standard techniques well known in the art. There is no detectable evidence of E. coli and the treatment is terminated.

EXAMPLE 43

Preparation and use of Triple Helix Probes

The sequences of microorganism genes required for proliferation of the present invention are scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches that could be used in triple-helix based strategies for inhibiting gene expression. Following identification of candidate homopyrimidine or homopurine stretches, their efficiency in inhibiting gene expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into a population of bacterial cells that normally express the target gene. The oligonucleotides may be prepared on an oligonucleotide synthesizer or they may be purchased commercially from a company specializing in custom oligonucleotide synthesis, such as GENSET, Paris, France.

The oligonucleotides can be introduced into the cells using a variety of methods known to those skilled in the art, including but not limited to calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection or native uptake.

Treated cells are monitored for a reduction in proliferation using techniques such as monitoring growth levels as compared to untreated cells using optical density measurements. The oligonucleotides that are effective in inhibiting gene expression in cultured cells can then be introduced *in vivo* using the techniques well known in that art at a dosage level shown to be effective.

In some embodiments, the natural (beta) anomers of the oligonucleotide units can be replaced with alpha anomers to render the oligonucleotide more resistant to nucleases. Further, an intercalating agent such as ethidium bromide, or the like, can be attached to the 3′ end of the alpha oligonucleotide to stabilize the triple helix. For information on the generation of oligonucleotides suitable for triple helix formation see Griffin et al. (Science 245:967-971 (1989), which is hereby incorporated by this reference).

EXAMPLE 44

Identification of Bacterial Strains from Isolated Specimens by PCR

Classical bacteriological methods for the detection of various bacterial species are time consuming and costly. These methods include growing the bacteria isolated from a subject in specialized media, cultivation on selective agar media, followed by a set of confirmation assays that can take from 8 to 10 days or longer to complete. Use of the identified sequences of the present invention provides a method to dramatically reduce the time necessary to detect and identify specific bacterial species present in a sample.

In one exemplary method, bacteria are grown in enriched media and DNA samples are isolated from specimens of, for example, blood, urine, stool, saliva or central nervous system fluid by conventional methods. A panel of PCR primers based on identified sequences unique to various species of microorganisms are then utilized in accordance with Example 12 to amplify DNA of approximately 100-200 nucleotides in length from the specimen. A separate PCR reaction is set up for each pair of PCR primers and after the PCR reaction is complete, the reaction mixtures are assayed for the presence of PCR product. The presence or absence of bacteria from the species to which the PCR primer pairs belong is determined by the presence or absence of a PCR product in the various test PCR reaction tubes.

Although the PCR reaction is used to assay the isolated sample for the presence of various bacterial species, other assays such as the Southern blot hybridization are also contemplated.

All documents cited herein are incorporated herein by reference in their entireties.